

Beiträge zur HOCHSCHULFORSCHUNG

2 | 2012

Höhle/Jacob/Teichler: Das Paradies nebenan

Lenhardt/Reisz/Stock: Überlebenschancen privater Hochschulen

Marx/Bornmann: Der Journal Impact Factor

Müller: Zitationen als Grundlage von Rankings

Beiträge zur HOCHSCHULFORSCHUNG

2 | 2012

Höhle/Jacob/Teichler: Das Paradies nebenan

Lenhardt/Reisz/Stock: Überlebenschancen privater Hochschulen

Marx/Bornmann: Der Journal Impact Factor

Müller: Zitationen als Grundlage von Rankings

Impressum

Beiträge zur Hochschulforschung

erscheinen viermal im Jahr

ISSN 0171-645X

Herausgeber: Bayerisches Staatsinstitut für Hochschulforschung
und Hochschulplanung, Prinzregentenstraße 24, 80538 München
Tel.: 089/2 1234-405, Fax: 089/2 1234-450

E-Mail: Sekretariat@ihf.bayern.de

Internet: <http://www.ihf.bayern.de>

Herausgeberbeirat:

Mdgt. a. D. Jürgen Großkreutz, Bayerisches Staatsministerium für Wissenschaft,
Forschung und Kunst, München

Dr. Lydia Hartwig, Bayerisches Staatsinstitut für Hochschulforschung und
Hochschulplanung, München

Professor Dr. Dorothea Jansen, Deutsche Hochschule für Verwaltungswissen-
schaften, Speyer

Professor Dr. Dr. h. c. Hans-Ulrich Küpper, Bayerisches Staatsinstitut für Hochschul-
forschung und Hochschulplanung und Ludwig-Maximilians-Universität, München

Thomas May, Wissenschaftsrat, Köln

Professor Rosalind Pritchard, AcSS, University of Ulster, United Kingdom

Redaktion: Dr. Lydia Hartwig (V.i.S.d.P.)

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung

E-Mail: Hartwig@ihf.bayern.de

Die abgedruckten Beiträge geben die Meinung der Verfasser wieder.

Graphische Gestaltung: Haak & Nakat, München

Satz: Dr. Ulrich Scharmer, München

Druck: Steinmeier, Deinungen

Ausrichtung, Themenspektrum und Zielgruppen

Die „Beiträge zur Hochschulforschung“ sind eine der führenden wissenschaftlichen Zeitschriften im Bereich der Hochschulforschung im deutschen Sprachraum. Sie zeichnen sich durch hohe Qualitätsstandards, ein breites Themenspektrum und eine große Reichweite aus. Kennzeichnend sind zudem die Verbindung von Wissenschaftlichkeit und Relevanz für die Praxis sowie die Vielfalt der Disziplinen und Zugänge. Dabei können die „Beiträge“ auf eine lange Tradition zurückblicken. Die Zeitschrift erscheint seit ihrer Gründung 1979 viermal im Jahr und publiziert Artikel zu Veränderungen in Universitäten, Fachhochschulen und anderen Einrichtungen des tertiären Bildungsbereichs sowie Entwicklungen in Hochschul- und Wissenschaftspolitik in nationaler und internationaler Perspektive.

Wichtige Themenbereiche sind:

- Strukturen der Hochschulen,
- Steuerung und Optimierung von Hochschulprozessen,
- Hochschulfinanzierung,
- Qualitätssicherung und Leistungsmessung,
- Studium und Studierende, Umsetzung des Bologna-Prozesses,
- Übergänge zwischen Schule, Hochschule und Arbeitsmarkt,
- Forschung und wissenschaftlicher Nachwuchs, akademische Karrieren,
- Frauen in Hochschulen und Wissenschaft,
- Wissenschaft und Wirtschaft,
- International vergleichende Hochschulforschung.

Die Zeitschrift veröffentlicht quantitative und qualitative empirische Analysen, Vergleichsstudien und Überblicksartikel, die ein anonymes Peer Review-Verfahren durchlaufen haben. Sie bietet die Möglichkeit zum Austausch von Forschungsergebnissen und stellt ein Forum für Hochschulforscher und Experten aus der Praxis dar. Zwei Ausgaben pro Jahr sind in der Regel einem aktuellen hochschulpolitischen Thema gewidmet, die beiden anderen sind inhaltlich nicht festgelegt. Es besteht die Möglichkeit, Aufsätze in deutscher und englischer Sprache einzureichen. Hinweise für Autoren befinden sich auf der letzten Seite.

Die „Beiträge“ richten sich an Wissenschaftler, die sich mit Fragen des Hochschulwesens und seiner Entwicklung befassen, aber auch an politische Entscheidungsträger, Hochschulleitungen, Mitarbeiter in Hochschulverwaltungen, Ministerien sowie Wissenschafts- und Hochschulorganisationen.

Inhalt

Editorial	4
Abstracts	6
Ester Höhle, Anna Katharina Jacob, Ulrich Teichler: Das Paradies nebenan? Zur Situation von Wissenschaftlerinnen und Wissenschaftlern an außer- universitären Forschungseinrichtungen und Universitäten in Deutschland	8
Gero Lenhardt, Robert D. Reisz, Manfred Stock: Überlebenschancen privater und öffentlicher Hochschulen im Ländervergleich	30
Werner Marx, Lutz Bornmann: Der Journal Impact Factor: Aussagekraft, Grenzen und Alternativen in der Forschungsevaluation	50
Harry Müller: Zitationen als Grundlage von Forschungsleistungsrankings – Konzeptionelle Überlegungen am Beispiel der Betriebswirtschaftslehre	68
Buchbesprechungen	94
Hinweise für Autoren	96

Editorial

Die „Beiträge zur Hochschulforschung“ nehmen wichtige Themen und Entwicklungen im Hochschulbereich aus unterschiedlichen Perspektiven in den Blick. Diese Ausgabe enthält Artikel, die sich mit der Arbeitssituation von Wissenschaftlern und Wissenschaftlerinnen in außeruniversitären Forschungseinrichtungen und Universitäten befassen, mit dem langfristigen Bestand privater und öffentlicher Hochschulen sowie mit der Messung von Forschungsleistungen. Die beiden Artikel zum Journal Impact Factor und zu Zitationen als Basis von Forschungsleistungsrankings leiten über zum nächsten Themenheft der „Beiträge zur Hochschulforschung“, das sich mit Qualitätssicherung in Lehre und Forschung beschäftigen wird.

Ester Höhle, Anna Katharina Jacob und Ulrich Teichler vergleichen in ihrem Artikel die Arbeitsbedingungen und Einstellungen von Wissenschaftlerinnen und Wissenschaftlern an Universitäten und außeruniversitären Forschungseinrichtungen. Sie zeigen, dass deren Arbeitssituation in manchen Bereichen große Unterschiede aufweist, so etwa beim Umfang der Befristung von Stellen für wissenschaftliche Mitarbeiterinnen und Mitarbeiter sowie bei der materiellen und personellen Ausstattung. Während die für Forschung aufgewandte Zeit in beiden Institutionen vergleichbar ist, zeigen sich die Beschäftigten an außeruniversitären Forschungseinrichtungen zufriedener mit ihrer Arbeitssituation und identifizieren sich tendenziell stärker mit ihrer Institution.

Der Artikel von Gero Lenhardt, Robert D. Reisz und Manfred Stock befasst sich mit der Wahrscheinlichkeit des Überlebens von privaten und öffentlichen Hochschulen in den USA, Chile und Deutschland. Die Autoren kommen zu dem Ergebnis, dass in allen drei Staaten Hochschulen mit erwerbswirtschaftlichem Charakter die geringsten Aussichten auf einen dauerhaften Bestand haben, etwas größer sind die der privaten gemeinnützigen und am größten die der öffentlichen Hochschulen.

Werner Marx und Lutz Bornmann diskutieren in ihrem Artikel die Grenzen der Anwendung des Journal Impact Factors als ein bibliometrischer Indikator zur Bemessung von Forschungsleistungen. Sie erinnern daran, dass dieser ursprünglich zur Einstufung von Zeitschriften innerhalb enger Fachkategorien entwickelt wurde und zeigen alternative, aussagefähigere Indikatoren für die quantitative Messung von Forschungsleistungen auf.

Einem ähnlichen Thema widmet sich der Artikel von Harry Müller „Zitationen als Grundlage von Forschungsleistungsrankings“. Er wirft am Beispiel der deutschsprachigen Betriebswirtschaftslehre die Frage auf, welche Probleme sich bei der Messung von Forschungsleistungen mit Hilfe von Zitationen ergeben. Der Autor

kommt zu dem Schluss, dass die Eignung einer Methode immer vor dem Hintergrund der Ziele des jeweils aufzustellenden Rankings beurteilt und die Ergebnisse nur innerhalb der vorgegebenen Grenzen interpretiert werden sollten.

Lydia Hartwig

Ester Höhle, Anna Katharina Jacob, Ulrich Teichler: The paradise next door? The situation of researchers in universities and research institutes outside the university sector

The article compares the situation of researchers in universities and research institutes outside the university sector in Germany, based on results of the international study "The Changing Academic Profession" of 2007. The analysis focuses on working conditions, attitudes and activities of the academics, grouped into university professors, directors of research institutes and research assistants at both types of institutions. It becomes obvious that the situation at both institutions varies greatly, in respect to e.g. fixed-term contracts for young researchers, resources, job satisfaction and productivity. Minor differences can be found in working contracts during the first years, time use and attitudes towards research. The comparison emphasises the different facets of science as a profession.

Gero Lenhardt, Robert D. Reisz, Manfred Stock: The chances of survival for private and public higher education institutions by comparing different states

The present paper studies the survival of higher education institutions in Germany, the USA and Chile by using a methodological approach of organisational sociology. The authors analyse the way in which different attributes of the organisation contribute to their survival by granting a special interest to the way the public or private affiliation influences survival. In all three countries the organisational survival rate of private for-profit higher education institutions is the lowest. The survival rate of private non-profit higher education institutions is somewhat higher; public higher education institutions have the highest survival rate. The authors discuss these results in the context of neo-liberal expectations that relate organisational fitness to criteria's of economic efficiency.

Werner Marx, Lutz Bornmann: The journal impact factor: relevance, limitations and alternatives in research evaluation

The journal impact factor, besides the h index, has become probably the best-known bibliometric indicator for the evaluation of research performance. Originally devised for the purposes of classifying journals as a whole within narrow subject categories, it is now frequently used to evaluate the publications of scientists and research institutes – though this is something that should not be equated with bibliometric analysis per se. This paper begins by discussing the limitations in the application of the journal impact factor with respect to its original purpose. The paper shows why it can be problematic even in the evaluation of journals. It goes on to demonstrate why journal

impact factors are unsuitable for evaluating the publications of individual scientists or entire research institutes. Finally, the paper points out alternative and more relevant metrics for quantitative (i. e. bibliometric) research evaluation. In conclusion, the paper fundamentally addresses the relevance and the limitations of bibliometrics in research evaluation.

**Harry Müller: Citations as a basis for research performance rankings.
Conceptual thoughts using the example of business studies**

This article addresses the problems and application fields of citations as a measure of research performance. First, it is pointed out why citations are a useful indicator of research performance and the limits of this approach. Thereon, the current methods for constructing research rankings are presented and critically evaluated. It turns out that for the German-based business studies, several characteristic features have to be considered: Research findings are often published in German, and apart from academic journals, monographs and editorships are equally important publication types. Taking this into account, it can be argued that citation based approaches are the most appropriate methodology for the assessment of research performance, despite some specific weaknesses.

Das Paradies nebenan? Zur Situation von Wissenschaftlerinnen und Wissenschaftlern an außeruniversitären Forschungseinrichtungen und Universitäten in Deutschland

Ester Höhle, Anna Katharina Jacob, Ulrich Teichler

In diesem Artikel werden auf Basis der internationalen Hochschullehrerstudie „The Changing Academic Profession“ von 2007 Universitäten und außeruniversitäre Forschungseinrichtungen in Deutschland verglichen. Im Zentrum der Analyse stehen die Arbeitsbedingungen, Einstellungen und Tätigkeiten der Wissenschaftlerinnen und Wissenschaftler¹, differenziert nach Universitätsprofessoren, Institutsdirektoren und wissenschaftlichen Mitarbeitern. Es zeigt sich, dass die Arbeitssituation von Wissenschaftlern in beiden Institutionen in einigen Bereichen große Unterschiede aufweist, wie z. B. der vertraglichen Befristung, der Ausstattung, der Berufszufriedenheit und der Produktivität. Dagegen sind die Unterschiede hinsichtlich der Vertragsbedingungen innerhalb der ersten Anstellungsjahre, der Zeitverwendung und der Einstellung zu Wissenschaft gering. Der Vergleich macht die verschiedenen Facetten von Wissenschaft als Beruf deutlich.

1 Zielsetzung und Anlage der Analyse

Deutschland wird international als das Land betrachtet, in dem das Konzept einer engen Verknüpfung von Forschung und Lehre als Leitidee für die moderne Universität entstanden ist. Die naheliegende Vermutung, dass in Deutschland eine Verzahnung von Forschung und Lehre mehr oder weniger flächendeckend besteht, lässt sich aber so nicht aufrechterhalten. Erstens gibt es seit Beginn der 1970er Jahre eine Zweigliederung der Hochschulsektoren, wobei die Fachhochschulen in größerem Umfang Lehraufgaben wahrnehmen (siehe *Enders 2010*). Zweitens zeigte die erste international vergleichende Hochschullehrerbefragung Anfang der 1990er Jahre, dass der wissenschaftliche Mittelbau an deutschen Hochschulen deutlich weniger Lehraufgaben hat als die Inhaber von Professuren (siehe *Enders und Teichler 1995*; vgl. auch *Altbach 1996*). Drittens gehört Deutschland zu den Ländern, in denen ein relativ großer Sektor öffentlich geförderter Forschung außerhalb der Universitäten angesiedelt ist, der Lehraufgaben nur in begrenztem Umfang übernimmt. Hier gibt es vielerlei personelle Verknüpfungen mit den Universitäten: Viele Institutsdirektoren sind zugleich

¹ Im Interesse der Lesbarkeit wird bei Personenbezeichnungen teilweise nur die männliche Form verwendet. Hier sind immer auch weibliche Personen gemeint, sofern kein entgegenstehender Hinweis erfolgt.

auf einer besonderen Professur („S-Professur“) an einer Universität tätig, und manche wissenschaftliche Mitarbeiter der Institute lehren an Universitäten. Nichtsdestotrotz sind die außeruniversitären Forschungsinstitute in erster Linie auf Forschung ausgerichtet. Wegen ihrer guten Ausstattung, ihrer internationalen Ausrichtung, ihrer Spitzenleistungen und ihrer guten Karriereperspektiven stehen die außeruniversitären Forschungseinrichtungen in dem Ruf, ihren Wissenschaftlern hervorragende Forschungsbedingungen zu bieten und ein Eldorado für aufstrebende Forscherinnen und Forscher zu sein, während die Forschungs- und Entwicklungsbedingungen für den wissenschaftlichen Nachwuchs an Universitäten in der Diskussion stehen. Diese Unterschiede sollen hier näher beleuchtet werden.

Die Institute genießen eine großzügigere Finanzierung als die Universitäten: 2007 betragen die Ausgaben an den vier Institutsgesellschaften 6,3 Milliarden Euro und an sämtlichen deutschen Universitäten 9,9 Milliarden Euro (*BMBF 2010, S. 41*). Der größte Teil der Gelder kommt den Naturwissenschaften zu Gute, während das kleinste Budget an die Geistes- und Sozialwissenschaften geht (*BMBF 2010, S. 479*). Die außeruniversitären Forschungsinstitute beschäftigten 2010 insgesamt 37.809 Personen als wissenschaftliches Personal (*Statistisches Bundesamt 2012, S. 53*), während an den Universitäten im gleichen Jahr 210.549 Personen als Wissenschaftler hauptberuflich tätig waren (*Statistisches Bundesamt 2011, S. 24*). Ähnlich wie jede Universität ein anderes Profil hat, haben auch die vier Institutsgesellschaften unterschiedliche Ausrichtungen: Während sich die 80 Institute der Max-Planck-Gesellschaft (MPG) insbesondere „auf freie Grundlagenforschung in innovativen Feldern“ konzentrieren, legt die Fraunhofer-Gesellschaft (FhG) den Fokus stärker auf die anwendungsorientierte Forschung. In ihren Instituten wird auch Forschung für die Industrie, Dienstleistungsunternehmen und die öffentliche Hand durchgeführt. Die 16 Forschungszentren der Helmholtz-Gemeinschaft (HGF) stellen Großgeräte und eine entsprechende Infrastruktur für nationale und internationale Forschungsgruppen aus dem naturwissenschaftlich-technischen und medizinischen Bereich bereit. Die 86 Zentren der Leibniz-Gemeinschaft betreiben in Kooperation mit Universitäten und Unternehmen vor allem „strategisch-programmatisch ausgerichtete Spitzenforschung“ in sechs thematisch orientierten Forschungsbereichen mit „einem Schwerpunkt auf nachfrageorientierter und interdisziplinärer Forschung“ (*BMBF 2010, S. 55; Hohn 2010, S. 460*).

Im Rahmen der international vergleichenden Hochschullehrerbefragung, die in den Jahren 2007 bzw. 2008 in 19 Ländern durchgeführt worden ist (vgl. *Jacob und Teichler 2011*), wurden in einigen Ländern – darunter auch Deutschland – ebenfalls die Angehörigen von Forschungsinstituten einbezogen (vgl. *Kogan und Teichler 2007; Locke und Teichler 2007*).

Für diesen Aufsatz wurde eine Sonderauswertung der Daten vorgenommen. Dabei ist zu bedenken, dass der Fragebogen mit Blick auf die an den Hochschulen Tätigen entwickelt worden ist. So sind manche Fragen für die an Forschungsinstituten Tätigen gar nicht oder nur in begrenztem Maße zutreffend. Umgekehrt fehlen Fragen, die ganz spezifisch auf die Situation in den Forschungsinstituten eingehen. An der Befragung haben insgesamt 465 Wissenschaftler an außeruniversitären Forschungseinrichtungen in Deutschland teilgenommen – etwa ein Drittel derjenigen, denen der Fragebogen zugesandt worden war – und zwar 136 Direktoren – das ist hier die vereinfachende Bezeichnung für Instituts- und Abteilungsleiter – und 329 wissenschaftliche Mitarbeiter:

- 45 bzw. 43 von der Max-Planck Gesellschaft (MPG),
- 14 bzw. 53 von der Fraunhofer Gesellschaft (FhG),
- 36 bzw. 85 von der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL) sowie
- 41 bzw. 148 von der Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF).

Von den wissenschaftlichen Mitarbeitern waren 78 noch nicht promoviert, und 239 promoviert. In der folgenden Analyse wurde keine Gewichtung nach unterschiedlichen Rücklaufquoten vorgenommen. Zu bedenken ist, dass einem höheren Anteil von Direktoren als von wissenschaftlichen Mitarbeitern der Fragebogen zugeschickt worden ist, um eine gewisse absolute Zahl von Antworten zu erreichen.

Die Aussagen der Befragten an den Forschungsinstituten wurden mit denen der etwa 150 Professoren und der über 800 wissenschaftlichen Mitarbeiter an den Universitäten in Deutschland verglichen. Die an Fachhochschulen in Deutschland Tätigen wurden hier nicht in den Vergleich einbezogen (zu deren Aussagen siehe *Jacob und Teichler 2011*).

2 Ausgewählte sozio-biographische Merkmale

Der Anteil der Frauen unter den Institutionsdirektoren beträgt neun Prozent. Er ist nur halb so hoch wie bei den Universitätsprofessoren. Unter den wissenschaftlichen Mitarbeitern an den Forschungsinstituten ist der entsprechende Anteil mit 21 Prozent zwar deutlich höher als bei den Direktoren, aber wiederum wesentlich geringer als an den Universitäten (38%). Auch bei Berücksichtigung der Fachrichtungen ergibt sich, dass der Anteil der Frauen in den Forschungseinrichtungen geringer ist als an den Universitäten:

- In den Geistes- und Sozialwissenschaften sind zehn Prozent der Direktorenposten von Frauen besetzt im Vergleich zu 24 Prozent der Universitätsprofessuren; die entsprechenden Anteile auf den Mitarbeiterpositionen sind 34 Prozent an Instituten gegenüber 54 Prozent an Universitäten.

- In den Natur- und Ingenieurwissenschaften sind unter den Antwortenden, die eine Direktorenposition an einem Forschungsinstitut innehaben, acht Prozent und somit weniger als der Anteil der Frauen in diesen Fächern, die eine Universitätsprofessur innehaben (13 %). Bei den wissenschaftlichen Mitarbeitern ergeben sich Werte von 19 Prozent an Instituten und 32 Prozent an Universitäten.

Das *Alter* der Institutsdirektoren beträgt im Durchschnitt 52 Jahre; das entspricht dem Durchschnittsalter von Professoren an Universitäten. Die wissenschaftlichen Mitarbeiter sind im Durchschnitt 45 Jahre alt und damit deutlich älter als an den Universitäten (37Jahre). Dies hängt, wie später aufgezeigt wird, damit zusammen, dass die wissenschaftlichen Mitarbeiter an den Instituten eher über die Qualifizierungsphase hinaus und dabei auch häufiger unbefristet beschäftigt werden. Im Hinblick auf die *familiäre Situation* ergeben sich jedoch zwischen den Angehörigen von Forschungsinstituten und Universitäten kaum Unterschiede (so z. B. bezüglich Familienstand, Berufstätigkeit des Partners, wissenschaftlicher Tätigkeit des Partners oder der Anzahl von Kindern im Haushalt).

Eine *ausländische Staatsgehörigkeit* ist bei den Institutsdirektoren häufiger der Fall als bei den Universitätsprofessoren: elf Prozent im Vergleich zu sieben Prozent; zur Zeit der Geburt waren sogar 16 Prozent bzw. neun Prozent Ausländer. Bei den wissenschaftlichen Mitarbeitern sind dagegen die Werte (7 % bzw. 6 % derzeit bzw. 9 % bzw. 8 % bei der Geburt) kaum unterschiedlich. Wie Tabelle 2 zeigt, ist der Ausländeranteil unter den Direktoren im Bereich der Geistes- und Sozialwissenschaften deutlich höher als unter den Universitätsprofessoren, während es bei den Natur- und Ingenieurwissenschaften kaum Unterschiede nach Institutionsart gibt.

Tabelle 1: Anteil der ausländischen Staatsangehörigen unter den Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007, nach Fachrichtungsgruppe (in Prozent)

	AUFE Prof			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
andere Staatsangehörigkeit bei Geburt	21	15	16	3	10	9	7	11	9	5	9	8
andere Staatsangehörigkeit bei erstem Abschluss	22	14	16	3	9	9	6	12	9	4	7	6
derzeit andere Staatsangehörigkeit	16	10	11	3	8	7	5	9	7	5	6	6
Anzahl (n)	18– 19	84– 86	102– 105	31– 32	214– 218	245– 250	52– 54	69– 71	121– 125	196– 198	538– 540	734– 738

Frage F9: Was war/ist Ihre Staatsangehörigkeit und das Land, in dem Sie tatsächlich lebten?

AUFE: außeruniversitäre Forschungsinstitute, Uni: Universität, Dir: Direktoren, Prof: Professoren, WiMi: wissenschaftliche Mitarbeiter, GS: Geistes- und Sozialwissenschaften, NI: Natur- und Ingenieurwissenschaften

3 Qualifikation und vorheriger Berufsweg

Das durchschnittliche *Alter zum Zeitpunkt der Promotion* lag bei den Institutsdirektoren im Durchschnitt bei 29 Jahren und bei den Professoren bei 30 Jahren. Die entsprechenden Werte für wissenschaftliche Mitarbeiter sind 31 bzw. 32 Jahre. Etwa sieben von zehn Institutsdirektoren sind *habilitiert*, gegenüber etwa acht von zehn Universitätsprofessoren. Das *Alter zum Zeitpunkt der Habilitation* betrug im Durchschnitt 37 bzw. 38 Jahre. Sofern die befragten wissenschaftlichen Mitarbeiter bereits habilitiert waren, lag das durchschnittliche Alter zum Zeitpunkt der Habilitation bei 41 bzw. 40 Jahren. Die geringfügigen Unterschiede erklären sich daraus, dass im Bereich der Natur- und Ingenieurwissenschaften, der in den außeruniversitären Forschungsinstituten stärker vertreten ist, die Promotion und Habilitation im Durchschnitt etwas früher erreicht wird als in den Geistes- und Sozialwissenschaften. Hinsichtlich der *beruflichen Mobilität* berichten Institutsdirektoren und Universitätsprofessoren etwa gleich häufig, dass sie seit ihrem höchsten Studienabschluss im Durchschnitt an 3,2 bzw. an 3,1 Institutionen tätig gewesen sind. Bei statistischer Kontrolle des Alters zeigt sich deutlich, dass wissenschaftliche Mitarbeiter an Universitäten häufiger mobil sind als ihre Kollegen an Forschungsinstituten (2,1 vs. 1,9); hier kommt zum Tragen, dass die Chance auf eine unbefristete Beschäftigung für wissenschaftliche Mitarbeiter an Forschungsinstituten höher ist als an Universitäten.

4 Beschäftigung und Einkommen

Eine *Teilzeitbeschäftigung* von Institutsdirektoren und Universitätsprofessoren ist eine so seltene Ausnahme, dass sich hier ein Rekurs auf Einzelheiten erübrigt. Dagegen sind 56 Prozent der nicht-promovierten wissenschaftlichen Mitarbeiter an Forschungsinstituten und 49 Prozent an Universitäten in den ersten sechs Jahren ihrer Beschäftigung Teilzeit beschäftigt. Nach einer Beschäftigungsdauer von mehr als sechs bzw. zwölf Jahren verringert sich diese Zahl besonders an Forschungsinstituten auch bei Nicht-Promovierten sehr deutlich, wie Tabelle 2 zeigt. Von den promovierten wissenschaftlichen Mitarbeitern an Forschungsinstituten mit mehr als zwölfjähriger Beschäftigungsdauer sind nur sechs Prozent Teilzeit beschäftigt.

Tabelle 2: Beschäftigungsumfang von wissenschaftlichen Mitarbeitern an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (in Prozent)

	AUFE nPr ≤6J	AUFE nPr >6J	AUFE Pr ≤12J	AUFE Pr >12J	Uni nPr ≤6J	Uni nP >6J	Uni Pr ≤12J	Uni Pr >12J
Vollzeit beschäftigt	42	83	82	93	46	62	81	69
Teilzeit beschäftigt	56	17	16	6	49	38	16	24
andere Vertragsformen	2	0	1	1	5	0	3	7
Anzahl (n)	48	30	74	163	290	76	278	218

Frage A8: In welchem Umfang sind Sie der an Ihrer Hochschule bzw. Forschungsinstitution beschäftigt?
 AUFE: außeruniversitäre Forschungsinstitute, Uni: Universität, Pr: promoviert, nPr: nicht promoviert

Befristet beschäftigt sind nur zwei Prozent der Institutsdirektoren und vier Prozent der Universitätsprofessoren. Dagegen ist bei wissenschaftlichen Mitarbeitern an Forschungsinstituten eine befristete Beschäftigung mit 43 Prozent sehr verbreitet und bei den wissenschaftlichen Mitarbeitern an Universitäten mit 79 Prozent überwiegend der Fall. Wiederum gibt es große Unterschiede nach dem Stadium des Berufswegs, wie Tabelle 3 belegt. Der höhere Anteil an Vollzeitbeschäftigung sowie der Umstand, dass nach sechsjähriger Beschäftigung ohne Promotion bzw. nach zwölfjähriger Beschäftigung mit Promotion an den Instituten der größte Teil unbefristet beschäftigt ist, erscheinen im Vergleich zu den Universitätsangestellten als ein nahezu paradiesischer Zustand.

Tabelle 3: Beschäftigungsdauer von wissenschaftlichen Mitarbeitern an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (in Prozent)

	AUFE				Uni			
	nPr ≤6J	nPr >6J	Pr ≤12J	Pr >12J	nPr ≤6J	nPr >6J	Pr ≤12J	Pr >12J
unbefristet beamtet	0	3	0	2	0	4	4	15
unbefristet angestellt	2	70	19	87	1	15	6	40
befristet, mit Aussicht auf unbefristeten Vertrag	0	3	22	4	4	9	4	6
befristet, ohne Aussicht auf unbefr. Vertrag	98	23	59	7	93	71	85	40
etwas anderes	0	0	0	1	2	0	1	1
Anzahl (n)	48	30	74	164	288	76	278	221

Frage A13: „Welche Beschäftigungsdauer ist bei Ihrem Beschäftigungsvertrag vorgesehen? (Bitte nur eine Angabe)“;

Unter den befragten Institutsdirektoren werden 81 Prozent entsprechend C4/W3-Professuren *eingestuft* bzw. vergütet und nur 19 Prozent entsprechend C3/W2. Bei den befragten Universitätsprofessoren betragen die entsprechenden Anteile 57 Prozent bzw. 43 Prozent. Das durchschnittliche *Bruttojahreseinkommen* aus der hauptberuflichen Tätigkeit liegt im Falle der Institutsdirektoren bei ca. 88.000 Euro und im Falle der Universitätsprofessoren bei ca. 72.000 Euro. In der oberen Besoldungsstufe liegen die Durchschnittswerte bei über 90.000 Euro einerseits und etwas unter 80.000 Euro andererseits. Allerdings berichten die Institutsdirektoren, dass sie im Durchschnitt nur 1.600 Euro *Nebeneinkünfte* im Jahr haben, die Universitätsprofessoren dagegen durchschnittlich 8.300 Euro. Bei gleicher Gehaltsstufe verbleibt für die Institutsdirektoren so lediglich ein Einkommensvorsprung von ca. 3.000 Euro. Das durchschnittliche Brutto-Einkommen für die hauptberufliche Tätigkeit liegt bei wissenschaftlichen Mitarbeitern an Forschungsinstituten mit etwa 41.000 Euro deutlich höher als an Universitäten mit fast 36.000 Euro. Über die Hälfte der Differenz ist darauf zurückzuführen, dass erstere seltener Teilzeit beschäftigt und im Durchschnitt älter sind; aber auch unter Berücksichtigung dieser beiden Faktoren zeigt sich, dass erstere im Schnitt in höhere Gehaltsklassen eingestuft sind. Die durchschnittlichen Nebeneinkünfte liegen bei den an Instituten Tätigen nach eigenen Angaben etwas unter 1.000 Euro und bei den an Universitäten Tätigen etwas über 1.000 Euro. Nach Kontrolle von Alter, Teilzeitbeschäftigung und Nebentätigkeit verbleibt ein Einkommensvorsprung von ca. 2.000 Euro für die wissenschaftlichen Mitarbeiter an den Forschungsinstituten. Die Angaben der Befragten deuten darauf hin, dass es zwischen den Dachorganisationen der Forschungsinstitute Unterschiede in der Vergütung gibt. Während bei den Direktoren die durchschnittlichen Vergütungen in der Max-Planck-Gesellschaft herausragen, haben umgekehrt die wissenschaftlichen Mitarbeiter an Max-Planck-Instituten im Durchschnitt das geringste Einkommen.

5 Arbeitssituation

Auch die *Bewertung der materiellen und personellen Ausstattung für die eigene wissenschaftliche Arbeit* fällt seitens der an außeruniversitären Instituten Tätigen weitaus positiver aus als seitens der an Universitäten Tätigen, wie Tabelle 4 belegt.

Tabelle 4: Bewertung der personellen und materiellen Ausstattung für die eigene Arbeit seitens der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland, nach Fachrichtungsgruppe (arithmetisches Mittel)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi.		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI.	Ges.
materielle Ausstattung												
technische Ausstattung für Lehre	2,1	2,4	2,4	2,0	2,1	2,1	2,7	2,7	2,7	2,7	2,5	2,6
Laboratorien	1,6	1,5	1,5	1,7	1,7	1,7	3,3	2,5	2,6	2,6	2,6	2,6
Forschungsausrüstung/ Instrumente	1,6	1,4	1,4	1,9	1,5	1,5	3,0	2,4	2,6	2,9	2,4	2,5
Computerausstattung	1,3	1,4	1,4	1,6	1,6	1,6	2,3	2,2	2,2	2,3	2,3	2,3
Bibliotheksausstattung	1,6	1,8	1,8	1,9	2,0	2,0	2,8	2,7	2,8	2,7	2,5	2,6
eigener Arbeitsraum	1,5	1,5	1,5	1,7	1,9	1,9	2,3	2,1	2,2	2,4	2,5	2,5
Telekommunikationstechnik	1,5	1,5	1,5	1,6	1,4	1,5	2,0	1,8	1,9	1,8	1,9	1,8
finanzielle Unterstützung der Forschung	1,6	1,9	1,9	2,1	2,2	2,2	3,8	3,3	3,5	3,6	3,3	3,4
personelle Ausstattung												
Sekretariatsunterstützung	1,8	1,9	1,9	2,1	2,3	2,3	2,8	2,6	2,7	2,6	2,9	2,8
Personal zur Unterstützung der Lehre	2,4	3,2	3,0	2,3	3,1	3,0	3,2	3,3	3,3	3,5	3,3	3,4
Anzahl (n)	21	92	113	38	242	280	56	73	129	207	557	765

Frage B3: Wie würden Sie die folgenden Einrichtungen, Ressourcen und Personal an Ihrer Institution bewerten, die Sie zur Unterstützung Ihrer Arbeit benötigen?

Antwortskala von 1= „Ausgezeichnet“ bis 5= „Sehr schlecht“

- Auf einer Skala von 1=„ausgezeichnet“ bis 5=„sehr schlecht“ bewerten die Institutsdirektoren die materielle Ausstattung in den sieben Kategorien im Durchschnitt mit 1,6 im Vergleich zu 2,7 seitens der Universitätsprofessoren. Sie bewerten zudem die personelle Ausstattung in zwei Kategorien im Durchschnitt mit 1,8 ein wenig zurückhaltender, aber der Unterschied zu den letzteren (3,0) ist noch größer.
- Die wissenschaftlichen Mitarbeiter an Forschungsinstituten bewerten die materielle Ausstattung im Durchschnitt mit 1,8 – also nur unwesentlich weniger positiv als die Direktoren. Der Abstand zur durchschnittlichen Bewertung seitens der wissenschaftlichen Mitarbeiter an Universitäten (2,6) ist etwas kleiner als im Falle der

Direktoren bzw. Professoren, aber ebenfalls bemerkenswert groß. Zur personellen Ausstattung äußern sich die wissenschaftlichen Mitarbeiter an den Instituten deutlich weniger positiv als die Direktoren: 2,4 im Durchschnitt, es bleibt aber auch ein nicht unbeträchtlicher Abstand zu Bewertungen der wissenschaftlichen Mitarbeiter an Universitäten: 3,0 im Durchschnitt.

Diese großen Unterschiede sind nicht darauf zurückzuführen, dass Natur- und Ingenieurwissenschaften in den Forschungsinstituten stärker vertreten sind, wie Tabelle 4 ebenfalls belegt. Im Bereich der Natur- und Ingenieurwissenschaften bewerten die Institutsdirektoren die materielle Ausstattung im Durchschnitt mit 1,6 im Vergleich zu 2,4 seitens der Universitätsprofessoren und die personelle Ausstattung im Durchschnitt mit 1,8 im Vergleich zu 2,8. Bei den wissenschaftlichen Mitarbeitern sind die Unterschiede mit 1,8 und 2,5 bzw. 2,2 und 3,0 nur geringfügig geringer. Im Bereich der Geistes- und Sozialwissenschaften bewerten die Institutsdirektoren die materielle Ausstattung im Durchschnitt mit 1,5 im Vergleich zu 2,8 seitens der Universitätsprofessoren und die personelle Ausstattung im Durchschnitt mit 1,6 im Vergleich zu 3,1. Bei den wissenschaftlichen Mitarbeitern sind die Unterschiede mit 1,9 und 2,6 bzw. 2,2 und 3,0 nur geringfügig geringer. Es zeigt sich lediglich, dass die Institutsdirektoren aus dem Bereich der Geistes- und Sozialwissenschaften ihre Ausstattung etwas positiver bewerten als ihre Kollegen aus dem Bereich der Natur- und Ingenieurwissenschaften; dagegen bewerten die Professoren an den Universitäten aus dem Bereich der Geistes- und Sozialwissenschaften ihre Ausstattung negativer als ihre Kollegen aus dem Bereich der Natur- und Ingenieurwissenschaften. Bei den wissenschaftlichen Mitarbeitern unterscheiden sich die Bewertungen kaum nach Fachrichtungsgruppen. Insgesamt zeichnen die an außeruniversitären Forschungseinrichtungen beschäftigten Wissenschaftler von ihren Arbeitsbedingungen ein deutlich günstigeres Bild als die Wissenschaftler an den Universitäten. Das betrifft insbesondere die Rahmenbedingungen für das Kerngebiet jener Einrichtungen – die Forschung.

6 Einstellungen zur Tätigkeit und berufliche Zufriedenheit

Angehörige der außeruniversitären Forschungsinstitute haben bei ihrer beruflichen Arbeit verständlicherweise eine eindeutige *Präferenz für Forschungstätigkeit*:

- 62 Prozent der Institutsdirektoren geben an, ihre Präferenz liege eindeutig in der Forschung, was nur bei 12 Prozent der Universitätsprofessoren der Fall ist. Dagegen ist hier mit 63 Prozent die häufigste Aussage, beides sei ihnen wichtig, jedoch mit einem stärkeren Akzent auf der Forschung; dies wurde von 37 Prozent der Institutsdirektoren angegeben, unter denen demzufolge eine Präferenz für die Lehre äußerst selten anzutreffen ist.

- Bei den wissenschaftlichen Mitarbeitern liegt die Präferenz noch häufiger eindeutig in der Forschung (70 %) als bei den Direktoren bzw. – und umso mehr – als bei ihren Kollegen an den Universitäten (33 %). Letztere haben somit häufiger eine eindeutige Präferenz in der Forschung als die Universitätsprofessoren.

Bemerkenswert ist, dass sich die an den außeruniversitären Instituten Tätigen *mit ihrer Institution* weitaus stärker *verbunden fühlen* als die an Universitäten Tätigen. Im Gefühl der Verbundenheit mit der Disziplin oder mit der wissenschaftlichen Arbeits-einheit (Fakultät, Fachbereich bzw. Abteilung) unterscheiden sich die befragten Gruppen jedoch weniger:

- 77 Prozent der Direktoren sagen, dass ihr Institut für sie wichtig ist, aber nur 46 Prozent der Universitätsprofessoren sagen dies über ihre Universität.
- Etwas geringer liegen die Werte bei den wissenschaftlichen Mitarbeitern auseinander, aber in der Richtung ergibt sich ein ähnlicher Unterschied nach Institutions-typ: 67 Prozent der wissenschaftlichen Mitarbeiter an Forschungsinstituten heben hervor, dass ihr Institut für sie wichtig ist, aber nur 40 Prozent der wissenschaftlichen Mitarbeiter an Universitäten sagen dies über ihre Universität.

In einer Frage nach den *Einstellungen zur beruflichen Situation* äußern sich die an den Forschungsinstituten Tätigen ähnlich zu den an Universitäten im Hinblick auf die Antwortvorgaben „Mein Beruf stellt *eine starke persönlich Belastung* dar“. Wie Tabelle 5 zeigt, stimmen dem fast die Hälfte der Institutsdirektoren und Professoren an Universitäten zu, dagegen nur knapp zwei Fünftel der wissenschaftlichen Mitarbeiter beider Institutionsarten. Dabei sind es eher die Befragten aus den Geistes- und Sozialwissenschaften, die im Beruf eine starke persönliche Belastung sehen.

Dagegen äußern sich die Direktoren und die wissenschaftlichen Mitarbeiter von Forschungsinstituten viel positiver als die Angehörigen der Universitäten im Hinblick auf eine andere Dimension der beruflichen Situation: Sie halten *die derzeitige Situation* seltener für *schlecht, eine wissenschaftliche Karriere im Fach der Befragten zu beginnen*. Dabei überrascht der Befund, dass die Befragten aus dem Bereich der Natur- und Ingenieurwissenschaften die Situation in dieser Hinsicht negativer einschätzen als die Befragten aus dem Bereich der Geistes- und Sozialwissenschaften.

Tabelle 5: Einstellungen der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland zu Aufgaben der wissenschaftlichen Tätigkeit 2007 (in Prozent*)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
schlechte Zeit**	18	21	20	35	28	29	56	28	42	51	41	43
nicht wieder***	9	8	8	11	13	13	19	17	14	17	19	18
Belastung****	36	48	47	22	39	37	46	51	48	33	36	36
Anzahl (n)	22	92	122*	37	242	297*	56	73	149*	207	555	828*

*Werte aus Gesamtauswertung

Vollständige Formulierungen im Fragebogen:

** Dies ist eine schlechte Zeit für junge Menschen, die eine wissenschaftliche Karriere in meinem Fach beginnen wollen.

*** Wenn ich noch einmal die Wahl hätte, würde ich nicht mehr Wissenschaftler/in an der Hochschule bzw. einem Forschungsinstitut werden. **** Mein Beruf stellt eine starke persönliche Belastung dar.

Frage B6: Wie ist Ihre Meinung zu den folgenden Äußerungen?

Antworten 1 und 2 auf einer Skala von 1 = „Stimme völlig zu“ bis 5 = „Stimme überhaupt nicht zu“

Bei der dritten zur beruflichen Situation gehörenden Frage – ob die Wissenschaftler sich erneut für eine wissenschaftliche Karriere entscheiden würden oder eher nicht – äußern sich die Institutsdirektoren positiver als die Universitätsprofessoren, während die Unterschiede zwischen den wissenschaftlichen Mitarbeitern der beiden Institutionsarten marginal sind. Auf die Antwortvorgabe „*Wenn ich noch einmal die Wahl hätte, würde ich nicht mehr Wissenschaftler/in an der Hochschule bzw. an einem Forschungsinstitut werden*“ äußern sich nur acht Prozent der Institutsdirektoren zustimmend, dagegen 14 Prozent der Universitätsprofessoren. Auch hier überrascht, dass sich die Befragten aus den Geistes- und Sozialwissenschaften unter den heutigen Bedingungen etwas stärker zu einer wissenschaftlichen Karriere bekennen als die aus den Natur- und Ingenieurwissenschaften. Nach den vorher genannten Befunden war es zu erwarten, dass die Institutsdirektoren in der Einschätzung der *allgemeinen beruflichen Zufriedenheit* sich zu einem höheren Anteil positiv äußern (93 %) als die Universitätsprofessoren (66 %) (siehe Tabelle 6). Die wissenschaftlichen Mitarbeiter an den Forschungsinstituten äußern sich sogar etwas häufiger zufrieden mit ihrer beruflichen Situation (77 %) als die Universitätsprofessoren und deutlich zufriedener als die wissenschaftlichen Mitarbeiter an Universitäten (55 %).

Tabelle 6: Zufriedenheit mit der beruflichen Situation insgesamt seitens der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (in Prozent*, arithmetisches Mittel)

	AUFE Dir	AUFE WiMi	AUFE Ges.	Uni Prof	Uni WiMi	Uni Ges.
1 sehr zufrieden	44	21	27	19	10	12
2	49	56	54	47	45	45
3	5	17	14	22	29	28
4	2	5	4	8	12	12
5 sehr unzufrieden	1	1	1	4	3	4
Anzahl (n)	121	298	419	86	828	1135**
Arithmetisches Mittel	1,7	2,1	2	2,5	2,3	2,5

Frage B7: Wie würden Sie insgesamt Ihre berufliche Zufriedenheit mit Ihrer derzeitigen Tätigkeit als Wissenschaftler einschätzen?

* Antworten auf einer Skala von 1 = „Sehr hoch“ bis 5 = „Sehr gering“

**Werte aus Gesamtauswertung

7 Einstellungen zur Wissenschaft und Akzente der eigenen Forschung

Wie schon erwähnt, haben die Beschäftigten der Forschungsinstitute die Möglichkeit, mehr Zeit mit Forschung zu verbringen als die Beschäftigten der Universitäten. Die Direktoren wenden 46 Prozent ihrer Zeit für Forschung auf, während es bei den Universitätsprofessoren 38 Prozent sind, die wissenschaftlichen Mitarbeitern an Instituten 63 Prozent, während es bei den wissenschaftlichen Mitarbeitern an Universitäten 53 Prozent sind. Obwohl bei den Institutsdirektoren Lehre nicht zu den Hauptaufgaben zählt, wenden sie dafür immerhin 12 Prozent ihrer Zeit auf, während Universitätsprofessoren 29 Prozent ihrer Zeit mit Lehraufgaben verbringen. Bei den wissenschaftlichen Mitarbeitern sind es zehn Prozent an den Instituten und 22 Prozent an den Universitäten. Die an außeruniversitären Forschungsinstituten Tätigen sind ebenso wie die Universitätsangehörigen gebeten worden, ihre *Einstellungen im Hinblick auf vier potenzielle Akzente der Wissenschaft* darzulegen: originäre Forschungsarbeit, Bilanzierung des Wissensstands, Anwendungsorientierung der Wissenschaft sowie Übertragung auf gesellschaftliche Problemlösungen. Dabei fallen die Aussagen zu den ersten beiden Themen nach Institutionsart sehr ähnlich aus, wie Tabelle 7 zeigt:

- 98 Prozent der Institutsdirektoren und 83 Prozent der Professoren an Universitäten halten originäre Forschungsarbeit für wichtig; die Werte sind im Falle der wissenschaftlichen Mitarbeiter mit 73 Prozent und 69 Prozent etwas niedriger.
- Nur 60 Prozent der Direktoren bzw. 61 Prozent der Professoren betonen, dass eine Bilanzierung des Wissensstands wichtig sei; bei den wissenschaftlichen Mitarbeitern sind die entsprechenden Werte mit 71 Prozent bzw. 67 Prozent etwas höher.

Dagegen ergeben sich große Unterschiede im Hinblick auf Anwendung und gesellschaftliche Problemlösung:

- Die Institutsdirektoren (48 %) befürworteten seltener als die Universitätsprofessoren (62 %) eine Anwendungsorientierung der Wissenschaften. In dieser Hinsicht unterscheiden sich allerdings die Aussagen der wissenschaftlichen Mitarbeiter kaum nach Institutionsart (66 % im Vergleich zu 67 %).
- Die Institutsdirektoren (46 %) sprechen sich seltener für eine Übertragung der Forschungsergebnisse auf praktische gesellschaftliche Problemlösungen aus als die Professoren an Universitäten (61 %). Bei den wissenschaftlichen Mitarbeitern ergibt sich ein Unterschied in der umgekehrten Richtung (51 % im Vergleich zu 44 %).

Tabelle 7: Einstellungen zur Wissenschaft seitens der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (in Prozent*, arithmetisches Mittel)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
originäre Forschungsarbeit**	86	90	89	76	72	73	85	88	83	70	69	69
Bilanzierung des Wissenstands***	55	62	60	65	71	71	56	65	61	63	68	67
Anwendungsorientierung****	45	49	48	65	67	66	62	63	62	66	67	67
gesellschaftliche Probleme*****	59	41	46	69	48	51	68	55	61	58	38	44
Anzahl (n)	22	92	122*	37	242	297*	56	73	149*	207	555	828*

Frage B5: Wie ist Ihre Meinung zu den folgenden Äußerungen?

*Antworten 1 und 2 auf einer Skala von 1 = „Stimme völlig zu“ bis 5 = „Stimme überhaupt nicht zu“

*Werte aus Gesamtauswertung

Vollständige Formulierungen im Fragebogen: ** Wissenschaft ist vor allem durch Vorbereitung und Durchführung originärer Forschungsarbeiten gekennzeichnet. *** Berichte zu erstellen, in denen die wichtigsten Trends und Ergebnisse der Forschung zusammengetragen werden, sind integraler Bestandteil der wissenschaftlichen Arbeit. **** Zu wissenschaftlicher Tätigkeit gehören Fragen der Anwendung im Alltag dazu“. ***** Wissenschaftler/innen in meinem Fachgebiet haben die berufliche Verpflichtung, ihr Wissen auf gesellschaftliche Probleme anzuwenden.

Dabei zeigen sich allerdings zwischen den Institutsdirektoren und den Universitätsprofessoren deutliche Unterschiede nach Fachrichtungsgruppen:

- An den Universitäten sind es eher die Vertreter der Geistes- und Sozialwissenschaften, die eine Verpflichtung sehen, Wissen auf gesellschaftliche Probleme zu übertragen, an den Instituten dagegen eher die Vertreter der Natur- und Ingenieurwissenschaften.

- Unter den Institutsdirektoren betonen die Vertreter der Natur- und Ingenieurwissenschaften häufiger als die Vertreter der Geistes- und Sozialwissenschaften, dass die Bilanzierung des Wissensstands wichtig ist, und auch, dass zur Wissenschaft Anwendung gehört. Dagegen gibt es bei den Universitätsprofessoren zu beiden Fragen kaum Unterschiede nach Fachrichtungsgruppe.

Lediglich in einer Hinsicht ergibt sich eine Gemeinsamkeit zwischen den Institutsdirektoren und den Universitätsprofessoren: In beiden Fällen betonen die Vertreter der Geistes- und Sozialwissenschaften häufiger als die Vertreter der Natur- und Ingenieurwissenschaften die Aufgabe der Wissenschaft, originäre Forschungsarbeiten zu leisten. Ähnlich klare Muster waren unter den wissenschaftlichen Mitarbeitern nicht anzutreffen. Im manchen, aber nicht allen Aspekten zu den allgemeinen Aussagen über die Aufgaben der Wissenschaft fallen die Äußerungen der Befragten anders aus, wenn sie die *Ausrichtung der eigenen Forschungsarbeit* charakterisieren:

- Die Institutsangehörigen bezeichnen ihre eigene Forschungsarbeit häufiger als „international bzw. international vergleichend“ orientiert als die an Universitäten tätigen Wissenschaftler: 84 Prozent der Direktoren sowie 68 Prozent der wissenschaftlichen Mitarbeiter an Instituten stehen 71 Prozent der Universitätsprofessoren und 47 der dortigen wissenschaftlichen Mitarbeiter mit entsprechenden Nennungen gegenüber.
- Die Universitätsprofessoren dagegen beschreiben ihre Tätigkeit etwas häufiger als „angewandt/praxisorientiert“ (64 % bzw. 58 %) und „gesellschaftlich orientiert/als Beitrag zur Verbesserung der Gesellschaft verstanden“ (48 % bzw. 41 %) als die Direktoren. Allerdings sehen die Institutsdirektoren häufiger ihre Arbeit „kommerziell ausgerichtet/auf Technologietransfer gerichtet“ (21 % bzw. 25 %). Dagegen sehen die wissenschaftlichen Mitarbeiter an den Instituten ihre Forschungstätigkeit häufiger als Beitrag zur Verbesserung der Gesellschaft als ihre Kollegen an den Universitäten (37 % bzw. 26 %).

Innerhalb der Forschungsinstitute ergeben sich im Durchschnitt große Unterschiede nach den vier Institutionsgruppen. In der Max-Planck-Gesellschaft wird Theorie- und Grundlagenorientierung betont sowie auf Internationalität bzw. internationalen Vergleich besonderer Wert gelegt; disziplinäre Orientierung liegt höher im Kurs als bei den anderen Instituten; dagegen wird weniger Wert auf Anwendungsorientierung, kommerzielle Verwendung der Ergebnisse und die Übertragung der Ergebnisse auf gesellschaftliche Probleme gelegt. Die Aussagen der Angehörigen von Fraunhofer-Instituten stehen dazu deutlich im Kontrast. Dazwischen liegen der Aussagen der Angehörigen der beiden anderen Gemeinschaften, wobei die an Leibniz-Instituten Tätigen eher in die Richtung der Max-Planck-Institute und die an Helmholtz-Instituten eher in die Richtung der Fraunhofer-Institute tendieren.

8 Forschungsaktivitäten

In der Beschreibung der *Forschungsaktivitäten* unterscheiden sich die Institutsdirektoren nicht erheblich von den Universitätsprofessoren (vgl. Tabelle 8):

- Die meisten Institutsdirektoren und etwas weniger Universitätsprofessoren (93 % vs. 83 %) haben innerhalb eines Jahres *Anträge zur Forschungsförderung* geschrieben. Immerhin über die Hälfte der wissenschaftlichen Mitarbeitern waren ebenfalls mit Antragstellung befasst (67 % vs. 57 %).
- In die *Vorbereitung* von Experimenten und anderen Untersuchungsarten waren über zwei Drittel der Direktoren und Professoren involviert (70 % vs. 64 %). Hier spielen die Mitarbeiter mit drei Vierteln eine große Rolle.
- Auch bei der *Beschaffung von Forschungsausrüstung* sind die Institutsdirektoren etwas häufiger aktiv als die Professoren (71 % vs. 61 %), das gleiche gilt für die wissenschaftlichen Mitarbeiter (63 % vs. 45 %).
- Ebenso sind die Direktoren bei der *Leitung von Forschungsteams und der Anleitung von Wissenschaftlern in der Forschungsarbeit* etwas häufiger involviert (97 % vs. 81 %) als Professoren. Auch hier spielen die wissenschaftlichen Mitarbeiter an Forschungsinstituten eine wichtige Rolle (70 % vs. 43 %).
- Ebenso geben die Direktoren und wissenschaftlichen Mitarbeiter an Instituten häufiger die *Verwaltung von Forschungsprojekten* an als die Kollegen an den Universitäten (89 % vs. 75 % und 61 % vs. 39 %).
- An der *Durchführung von Experimenten und Erhebungen* sind die wissenschaftlichen Mitarbeitern naheliegenderweise häufiger beteiligt als die Professoren bzw. Direktoren (52 % der Direktoren und 71 % der wissenschaftlichen Mitarbeiter). Hier gibt es kaum Unterschiede zwischen den Institutionen.
- Mit dem *Schreiben von Forschungsberichten und Publikationen* sind mehr als neun Zehntel aller Befragten befasst.

Die genannten Unterschiede hängen weitgehend mit der Projektförderung und der Zugehörigkeit zu Fachrichtungsgruppen zusammen. Wie Tabelle 8 zeigt, sind jedoch auch innerhalb der beiden Fachrichtungsgruppen in geringerem Umfang Unterschiede zwischen den Angehörigen von Forschungsinstituten und Universitäten erkennbar.

Tabelle 8: Aktuelle Forschungstätigkeiten der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007, nach Fachrichtungsgruppe (in Prozent, Mehrfachantworten möglich)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
Verfassen von Anträgen zur Forschungsförderung	86	97	93	74	67	67	89	78	83	54	57	57
Vorbereiten von Experimenten, Erhebungen	52	75	70	53	78	74	74	44	64	55	82	75
Beschaffung von Forschungsausrüstung	29	83	71	15	70	63	79	35	61	29	49	45
Leitung eines Forschungsteams	95	97	97	62	71	70	97	63	81	23	49	43
Verwaltung von Forschungsprojekten	86	90	89	65	60	61	88	63	75	29	41	39
Durchführung von Experimenten usw.	52	53	52	53	75	71	58	38	51	51	83	75
Schreiben von Forschungsberichten und Publikationen	95	98	97	100	91	92	95	91	93	89	92	91
Involvierung in Technologietransfer	5	43	37	6	36	32	37	7	24	4	15	13
Anzahl (n)	22	106	136	42	265	329	74	59	152	226	587	888

Frage D3: Haben Sie im Rahmen Ihrer aktuellen Forschungstätigkeiten innerhalb eines Jahres folgende Tätigkeiten unternommen?
Antwortkategorien „ja“ und „nein“.

Für Institutsdirektoren sind *Forschungskooperationen* häufiger die gängige Praxis als für Universitätsprofessoren; ähnliches gilt für wissenschaftliche Mitarbeiter. Ebenso lässt sich sagen, dass im Bereich der Natur- und Ingenieurwissenschaften Forschungs-kooperationen stärker verbreitet sind als im Bereich der Geistes- und Sozialwissenschaften (siehe Tabelle 9):

- Von den Direktoren aus dem Bereich der Natur- und Ingenieurwissenschaften geben jeweils über 90 Prozent an, dass sie *mit Personen an ihrer Institution, an anderen Institutionen in Deutschland und in anderen Ländern* kooperieren; die Werte seitens der Professoren an Universitäten stehen dem kaum nach. Dementsprechend geben nur 31 Prozent der Institutsdirektoren in dieser Fachrichtungsgruppe im Unterschied zu 57 Prozent der Professoren an Universitäten an, mindestens in einem Projekt *allein* zu arbeiten.
- Wissenschaftliche Mitarbeiter an Forschungsinstituten im Bereich der Natur- und Ingenieurwissenschaften sind nach eigenen Angaben etwas weniger häufig in Forschungs-kooperationen eingebunden als Direktoren, Forschungs-kooperationen sind aber bei ihnen deutlich stärker verbreitet als bei den wissenschaftlichen Mit-

arbeitern ihrer Fachrichtungsgruppe an den Universitäten. Dagegen führen wissenschaftliche Mitarbeiter der Natur- und Ingenieurwissenschaften an Universitäten häufiger auch alleine Forschungsprojekte durch als ihre Kollegen an den Forschungsinstituten (61 % bzw. 41 %).

Tabelle 9: Zusammenarbeit mit anderen Wissenschaftlern seitens der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (in Prozent, Mehrfachantworten möglich)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
Sie arbeiten allein an mindestens einem Projekt	62	31	37	50	41	42	81	57	67	80	61	66
Sie kooperieren in der Forschung mit Personen an Ihrer Institution	100	92	94	76	87	85	61	87	76	52	73	67
Sie kooperieren in der Forschung mit Personen an anderen Institutionen in Deutschland	86	92	91	71	84	82	74	93	85	44	59	55
Sie arbeiten in der Forschung mit Personen zusammen, die in anderen Ländern tätig sind	100	94	95	65	76	74	58	88	75	33	43	41
Anzahl (n)	21	89	110	34	232	266	56	73	130	194	543	738

Frage D1: Wie würden Sie Ihre aktuellen Forschungsvorhaben beschreiben?

- Direktoren aus dem Bereich der Geistes- und Sozialwissenschaften arbeiten ähnlich häufig wie ihre Kollegen aus den Natur- und Ingenieurwissenschaften in Forschungs Kooperationen. Dagegen sind die entsprechenden Werte bei den Universitätsprofessoren der Natur- und Ingenieurwissenschaften geringer: 61 Prozent kooperieren an ihrer Institution, 74 Prozent im Inland und 58 Prozent im Ausland. Die Direktoren der Geistes- und Sozialwissenschaften unterscheiden sich aber deutlich von ihren Kollegen der Natur- und Ingenieurwissenschaften darin, dass viele von ihnen ihre Projekte alleine durchführen – wenn auch nicht so häufig wie die Universitätsprofessoren in den Geistes- und Sozialwissenschaften (62 % bzw. 81 %).
- Die wissenschaftlichen Mitarbeiter im Bereich der Geistes- und Sozialwissenschaften an Forschungsinstituten sind auch seltener in Forschungs Kooperationen eingebunden als Direktoren; Forschungs Kooperationen sind bei ihnen aber dennoch deutlich stärker verbreitet ist als bei den wissenschaftlichen Mitarbeitern ihrer Fachrichtungsgruppe an den Universitäten: Innerhalb ihrer eigenen Institution kooperieren 76 Prozent im Vergleich zu 52 Prozent, im Inland 71 Prozent bzw. 44 Prozent und im Ausland 65 Prozent bzw. 33 Prozent. Dagegen führen wissenschaftliche Mitarbeiter dieser Fachrichtungsgruppe an Universitäten häufiger auch Forschungsprojekte alleine durch als ihre Kollegen an den Forschungsinstituten (81 % bzw. 50 %).

Die befragten Institutsdirektoren sind häufiger Mitglied einer Evaluations- oder Gutachterkommission als die Professoren an den Universitäten (80 % gegenüber 63 %). Fast zwei Drittel der an Instituten Befragten teilten weiterhin mit, in wissenschaftlichen Kommissionen auf Landes-, Bundes- oder internationaler Ebene tätig zu sein (62 % vs. 42 %). Dagegen sind fast ein Fünftel der Universitätsprofessoren gesellschaftspolitisch aktiv (in der Politik, gemeinnützigen sozialen Organisationen, Gewerkschaften u.a.), aber weniger als ein Zehntel der Institutsdirektoren. Wissenschaftliche Mitarbeiter üben derartige Funktionen seltener aus, aber es zeigt sich ein ähnliches Muster der Unterschiede zwischen Forschungsinstituten und Universitäten.

9 Veröffentlichungen

Hinsichtlich der *Publikationstätigkeit* sind die Institute den Universitäten gegenüber im Vorteil. Das gilt eindeutig für die Zahl der *Artikel in Zeitschriften und Sammelbänden*, die innerhalb der drei letzten Jahre veröffentlicht worden sind: Hier nennen im Bereich der Natur- und Ingenieurwissenschaften die Direktoren 24,1 Artikel, die Professoren dagegen 17,2 Artikel (siehe Tabelle 10). Jeweils drei Artikel weniger (21,2 und 14,3) werden aus dem Bereich der Geistes- und Sozialwissenschaften berichtet. Bei den wissenschaftlichen Mitarbeitern sind die entsprechenden Werte in den Natur- und Ingenieurwissenschaften 8,4 und 4,6 sowie in den Geistes- und Sozialwissenschaften 8,2 und 5,3. Auch bei anderen Textformen sind die Werte bei den Angehörigen der Forschungsinstitute deutlich höher: Das gilt für *Forschungsberichte* wie für *Papiere, die auf Tagungen vorgelegt worden sind*. Weniger konsistent sind die Unterschiede bei der *Autorenschaft von Büchern*, bei der *Herausgabe von Büchern* und bei der *Publikation von Fachartikeln in Tageszeitungen und Zeitschriften*.

Tabelle 10: Anzahl der innerhalb von drei Jahren abgeschlossenen wissenschaftlichen Werke seitens der Befragten an außeruniversitären Forschungsinstituten und Universitäten in Deutschland 2007 (arithmetisches Mittel)

	AUFE Dir			AUFE WiMi			Uni Prof			Uni WiMi		
	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.	GS	NI	Ges.
wiss. Bücher als Autor	0,9	0,3	0,5	0,9	0,5	0,5	0,9	0,5	0,7	0,4	0,2	0,2
wiss. Bücher als Herausgeber	2,4	0,5	0,9	0,6	0,2	0,2	2,2	0,6	1,2	0,4	0,1	0,2
Beiträge für wiss. Bücher bzw. Zeitschriften	21,2	24,3	23,4	8,2	8,8	8,6	14,5	17,2	15,4	5,3	4,6	4,8
Forschungsberichte für ein gefördertes Projekt	2,5	7,3	6,2	2,8	3,7	3,5	1,5	4,7	3,1	1,1	1,6	1,5
auf wiss. Konferenzen/Tagungen vorgelegte Papiere	19,2	17,3	17,7	6,1	8,4	8,0	6,1	14,9	10,5	3,7	5,2	4,8
für Tageszeit./Zeitschriften geschriebene Fachartikel	2,1	4,1	3,6	4,0	1,5	1,8	2,7	2,7	2,7	1,1	0,9	1,0
Patentierungen von Verfahren oder Erfindungen	0	1,2	1,0	0	0,6	0,5	0,1	1,0	0,6	0	0,2	0,1
Anzahl (n)	20	87	107	33	223	256	54	73	127	182	514	696

Frage D4: Wie viele der folgenden wissenschaftlichen bzw. künstlerischen Publikationen oder andere Beiträge haben Sie in den letzten drei Jahren abgeschlossen?

In einigen Publikationen der international vergleichenden Hochschullehrerstudie wurde ein *Indexmaß für die Veröffentlichungshäufigkeit* verwendet. Dabei wurden Buchautorenschaften und -herausgaben jeweils mit drei Punkten, Aufsätze und Forschungsberichte jeweils mit zwei Punkten sowie Konferenzpapiere und Beiträge zu Tageszeitungen und allgemeinen Zeitschriften mit einem Punkt eingestuft. Demnach ergibt sich bei den Institutsdirektoren ein Indexwert von 80 und bei den Universitätsprofessoren von 55. Bei den wissenschaftlichen Mitarbeitern betragen die entsprechenden Werte 35 und 20. Dieser deutliche institutionelle Vorsprung für die außeruniversitären Institute ist jedoch zu relativieren: Berücksichtigt man, dass Publikationen im Bereich der Natur- und Ingenieurwissenschaften etwas häufiger sind, dass die Institutsdirektoren ein Drittel mehr Zeit für die Forschung aufbringen können als die Universitätsprofessoren und dass bei den wissenschaftlichen Mitarbeitern die Publikationshäufigkeit mit der Qualifikationsstufe und einer Tätigkeit über die Qualifikationsphase hinaus steigt, so ist an den außeruniversitären Instituten keine höhere „Produktivität“ festzustellen.

10 Die Ergebnisse im Überblick

Die befragten Institutsdirektoren sind – so lässt sich zur Beschäftigungssituation feststellen – meist wie Universitätsprofessoren der oberen Kategorie (C4/W3) eingestuft, während bei Professoren etwa 40 Prozent der C3/W2-Kategorie angehören. Innerhalb

der gleichen Kategorie sind die Einkommen aus Haupt- und Nebentätigkeit für die Institutsdirektoren nach eigenen Angaben nur 3 bis 4 Prozent höher als bei den Universitätsprofessoren. Ähnlich gering sind die Unterschiede bei den wissenschaftlichen Mitarbeitern der etwa gleichen Altersgruppe. Wissenschaftliche Mitarbeiter an den Forschungsinstituten haben jedoch offenkundig eine größere Chance als solche an Universitäten, an ihrer Institution länger tätig zu sein und einen unbefristeten Vertrag zu erhalten – ein wesentlicher Aspekt, der die Institute gegenüber den Universitäten attraktiver macht, da dort ein hoher Anteil der wissenschaftlichen Mitarbeiter auch nach der Promotion befristet beschäftigt ist.

Obwohl die Aufgabe der Institute in der Forschung und die der Universitäten in Forschung und Lehre liegen, gibt es keinen sehr großen Unterschied im Anteil der für Forschung aufgewandten Arbeitszeit. Dagegen bewerten die Angehörigen der Forschungsinstitute die materiellen und personellen Arbeitsbedingungen weitaus positiver als ihre Kollegen an den Universitäten. Diese Differenz ist im Bereich der Geistes- und Sozialwissenschaften sogar noch größer als im Bereich der Natur- und Ingenieurwissenschaften. So kann es auch nicht überraschen festzustellen, dass die an außeruniversitären Forschungsinstitutionen Tätigen eine höhere berufliche Zufriedenheit haben als ihre Kollegen an Universitäten und sich stärker mit ihrer Institution identifizieren.

Fast alle Institutsdirektoren und fast alle Universitätsprofessoren in den Natur- und Ingenieurwissenschaften berichten, dass sie in der Forschung mit anderen kooperieren; nicht ganz so häufig ist das bei den Professoren in den Geistes- und Sozialwissenschaften der Fall. Wissenschaftliche Mitarbeiter von Forschungsinstitutionen sind stärker als solche von Universitäten in Kooperationen eingebunden. Daneben sind aber bemerkenswerte Anteile der Befragten in Projekten als Einzelne tätig; nur bei den Direktoren aus dem Bereich der Natur- und Ingenieurwissenschaften kommt dies kaum vor. Die Einstellungen zur Forschung und die eigenen Akzente der Tätigkeit lassen sich nicht durchgängig kontrastiv zwischen den Forschungsinstituten und den Universitäten einordnen. Im Durchschnitt betonen die Institutsdirektoren jedoch die praktische und gesellschaftliche Relevanz der Forschung weniger als die Universitätsprofessoren. In vieler Hinsicht, besonders aber bei der Herkunft der Direktoren im Vergleich zu den Universitätsprofessoren, ist bei den außeruniversitären Forschungsinstituten eine höhere Internationalität erkennbar als bei den Universitäten.

Die Zahl der Veröffentlichungen ist bei Institutsdirektoren im Durchschnitt deutlich höher als bei Universitätsprofessoren und unterscheidet sich auch bei den wissenschaftlichen Mitarbeitern an Forschungsinstituten deutlich von denen an Universitäten. Das ist für die Erfolgsbilanzen der Institutionen ein erheblicher Unterschied. Berücksichtigen wir dagegen, dass den Universitätsangehörigen weniger Zeit für Forschung

zur Verfügung steht und mehr wissenschaftliche Mitarbeiter an Universitäten sich in einer frühen Qualifizierungsphase befinden, so ist – trotz der günstigeren Ressourcen für die Forschung an den Forschungsinstituten – keine höhere relative „Produktivität“ der Wissenschaftler an den außeruniversitären Forschungsinstituten zu erkennen. Ein Paradies nebenan? Bei allen Unterschieden – insbesondere in der Ausstattung sowie den Vertragsbedingungen für wissenschaftliche Mitarbeiter – sind die Differenzen hinsichtlich des Tätigkeitsprofils vielfach geringer ausgefallen als man vielleicht hätte erwarten können.

Literatur

Altbach, Philip G. (Hg.) (1996): The International Academic Profession: Portraits of 14 Countries. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

Bundesministerium für Bildung und Forschung (2010): Bundesbericht Forschung und Innovation 2010. Bonn und Berlin: Bundesministerium für Bildung und Forschung.

Enders, Jürgen (2010): Hochschulen und Fachhochschulen. In: Simon, Dagmar/Knie, Andreas/Hornbostel, Stefan (Hg.): Handbuch Wissenschaftspolitik. Wiesbaden: VS-Verlag für Sozialwissenschaften, S. 443–456.

Enders, Jürgen/Teichler, Ulrich (1995): Berufsbild der Lehrenden und Forschenden an Hochschulen. Bonn: Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie.

Hohn, Hans-Willy (2010): Außeruniversitäre Forschungseinrichtungen. In: Simon, Dagmar/Knie, Andreas/Hornbostel, Stefan (Hg.): Handbuch Wissenschaftspolitik. Wiesbaden: VS-Verlag für Sozialwissenschaften, S. 457–477.

Jacob, Anna Katharina/Teichler, Ulrich (2011): Der Wandel des Hochschullehrerberufs im internationalen Vergleich: Ergebnisse einer Befragung in den Jahren 2007/08. Bonn und Berlin: Bundesministerium für Bildung und Forschung.

Kogan, Maurice/Teichler, Ulrich (Hg.) (2007): Key Challenges to the Academic Profession. Kassel: Internationales Zentrum für Hochschulforschung. Werkstattberichte, Nr. 65).

Locke, William/Teichler, Ulrich (Hg.) (2007): The Changing Conditions for Academic Work and Careers in Select Countries. Kassel: Internationales Zentrum für Hochschulforschung Kassel. Werkstattberichte, Nr. 66).

Statistisches Bundesamt (2011): Bildung und Kultur. Personal an Hochschulen 2010. Fachserie 11 Reihe 4.4. Wiesbaden.

Statistisches Bundesamt (2012): Finanzen und Steuern. Ausgaben, Einnahmen und Personal der öffentlichen und öffentlich geförderten Einrichtungen für Wissenschaft, Forschung und Entwicklung, Berichtszeitraum 2010. Fachserie 14 Reihe 3.6. Wiesbaden.

Anschriften der Autoren:

Ester Ava Höhle, M.A.
Prof. Dr. Dr. h. c. Ulrich Teichler
Internationales Zentrum für Hochschulforschung
INCHER-Kassel
Universität Kassel
Mönchebergstr. 17
34109 Kassel
E-Mail:
hoehle@incher.uni-kassel.de
teichler@incher.uni-kassel.de

Dr. Anna Katharina Jacob
Zentrum für Hochschul- und Qualitätsentwicklung (ZfH)
Universität Duisburg-Essen
Keetmanstraße 3–9
47058 Duisburg
E-Mail: katharina.jacob@uni-due.de

Ester Ava Höhle ist wissenschaftliche Mitarbeiterin am Internationalen Zentrum für Hochschulforschung der Universität Kassel.

Ulrich Teichler ist Professor für Berufs- und Hochschulforschung an der Universität Kassel und war langjähriger Geschäftsführender Direktor des Internationalen Zentrums für Hochschulforschung.

Anna Katharina Jacob ist Akademische Rätin am Zentrum für Hochschul- und Qualitätsentwicklung der Universität Duisburg-Essen.

Überlebenschancen privater und öffentlicher Hochschulen im Ländervergleich

Gero Lenhardt, Robert D. Reisz, Manfred Stock

Der Beitrag untersucht die Überlebenschancen von Hochschulen in Deutschland, den USA und Chile auf der Basis der soziologischen Organisationsökologie. Er analysiert, auf welche Weise verschiedene Organisationsmerkmale das Überleben von Hochschulen beeinflussen. Dabei ist der Vergleich von Überlebenswahrscheinlichkeiten in Abhängigkeit von der Trägerschaft, also im Vergleich von öffentlichen und privaten Hochschulen, von besonderem Interesse. In allen drei Ländern, so zeigt sich, ist die Überlebenswahrscheinlichkeit der privaten Hochschulen mit erwerbswirtschaftlichem Charakter am geringsten. Etwas größer ist die der privaten gemeinnützigen und am größten die der öffentlichen Hochschulen. Diese Befunde werden vor dem Hintergrund von neoliberalen Erwartungen diskutiert, die das Überleben von Hochschulen mit Wirtschaftlichkeitskriterien in Zusammenhang bringen.

1 Fragestellung

Neben der katholischen Kirche und der Bank von England zählt die Universität zu den ältesten Organisationen des Abendlandes. Dass die einzelnen Universitäten und anderen Hochschulen überleben, versteht sich aber keineswegs von selbst. Viele von ihnen sind im Laufe der Geschichte verschwunden. So wurde in Deutschland nach 1800 ungefähr jede zweite Universität geschlossen, als sich ihre Träger, die konfessionellen Kleinstaaten, auflösten (*Riddle 1996*).

In jüngster Zeit wird das Überleben von Hochschulen und Universitäten vor allem mit Wirtschaftlichkeitskriterien in Zusammenhang gebracht. Eine Orientierung an markt- und betriebswirtschaftlichen Regelungen, so die These neoliberaler Beobachter der gegenwärtigen Hochschulentwicklung, sei für die Überlebensfähigkeit entscheidend. Hochschulen mit einem privaten Träger seien dazu sehr viel besser in der Lage als die öffentlichen. Sie könnten ihr Bildungsangebot auf Nischen spezialisieren, sich flexibler an den rapiden gesellschaftlichen Wandel anpassen, insbesondere an den Wechsel der Bildungsnachfrage, und sie könnten diese effizienter bedienen. All dies verschaffe ihnen Vorteile gegenüber den öffentlichen Hochschulen.

So begründet die Beratungsfirma SSC Consult in Köln die besondere Attraktivität betriebs- und marktwirtschaftlich operierender Einrichtungen wie folgt:

„Neben der Tatsache, dass man als Investor einem so genannten Meta-Trend folgt, sind auch die Geschäftsmodelle der privaten Hochschulen unter Renditegesichtspunkten interessant weil skalierbar. Sind die Kosten erst einmal gedeckt, bringt jeder neue Student Gewinn. Daher beruhen die nahe liegenden Wertsteigerungsstrategien zu großen Teilen auch auf der Skalierbarkeit des Geschäftsmodells. Auch die Planbarkeit zeichnet sich im Vergleich zu vielen anderen Geschäftsmodellen als überaus stabil aus: haben sich die Studierenden in einem Studiengang erst einmal eingeschrieben, stehen die zukünftigen Cash Flows bereits fest. Die Abbruchraten sind insbesondere im Vergleich zu staatlichen Institutionen überaus gering“ (Savelberg 2008, S. 184 f.).

Es trifft gewiss zu, dass zusätzliche Studierende privaten Hochschulen zusätzliche Einnahmen verschaffen, während sie in den öffentlichen vor allem als eine zusätzliche Belastung angesehen werden. Aber ob der „Cash Flow“ so sicher rinnt, dass er den Bestand einer Hochschule sichert, ist ungewiss. Ebenso steht dahin, wie die privaten Hochschulen jenen nicht-monetären gesellschaftlichen Erfordernissen genügen, die die Wirtschaftstheorie als „Datenkranz“ bezeichnet, an den Rand ihrer Aufmerksamkeit rückt oder gänzlich aus dem Auge verliert. Wenn beispielsweise in der Gesellschaft die Verknüpfung von Lernen, Lehre und Forschung als vorbildlich gilt, dann können private Hochschulen schnell in eine Krise geraten, wenn sie aus Gründen betriebswirtschaftlicher Effizienz auf Forschung verzichten und damit an Reputation und Nachfrage seitens der Studierenden verlieren.

Erwägungen wie diese lassen die Frage entstehen, ob Hochschulen mit privatem Träger tatsächlich besonders überlebenstüchtig sind, überlebenstüchtiger jedenfalls als die öffentlichen. Im Folgenden soll dieser Frage im Rückgriff auf den Untersuchungsansatz der *organizational ecology* nachgegangen werden, der maßgeblich von Hannan und Freeman (*Hannan/Freeman 1989; Hannan/Carrol 1992*) entwickelt wurde. Analyseeinheiten sind dabei nicht einzelne Organisationen, sondern Organisationspopulationen einer Vielzahl von Einzelorganisationen.

Obwohl Hannan und Freeman in ihren organisationsökologischen Analysen sich nicht mit dem Überleben von Hochschulen befasst haben, sind sie auf einen Befund gestoßen, der sich auf unsere Frage beziehen lässt. Sie haben die Überlebensfähigkeit von generalistischen und spezialisierten Organisationsformen in Abhängigkeit von der Dynamik ihrer Umwelten untersucht. Die Autoren zeigen, dass in Relation zu Umwelten, die sich relativ rasch und stark ändern – so genannte „fine-grained high variability“ Umwelten¹ – (*Hannan/Freeman 1989, S. 322*), die Überlebenswahrscheinlichkeit von spezialisierten Organisationsformen größer ist als jene von generalistischen. Als

¹ „Fine grained“ heißt, dass die Dauer einer stabilen Phase der Umwelt kurz ist im Vergleich zur typischen Lebensdauer der Organisation. „High variability“ heißt, dass die Änderungen der Umwelt von den bestehenden Strukturen der Organisation nicht mehr aufgefangen werden können.

ein Grund dafür wird genannt, dass bei rasch wechselnden Umweltzuständen die Kosten für strukturelle Anpassungsmaßnahmen bei den generalistischen Organisationen sehr groß seien. Demgegenüber erscheine es als vorteilhafter, eine bestimmte spezialisierte Struktur gegenüber den raschen Zustandswechseln der Umwelt durchzuhalten. Überträgt man dies auf Hochschulorganisationen und geht mit den vorherrschenden Diagnosen der Hochschulforschung (*Skilbeck 1997; Wächter 2004*) von der Voraussetzung aus, dass sich deren Umwelten rasch wandeln, so müssten spezialisierte Einrichtungen größere Überlebenschancen haben. Im Vergleich zu öffentlichen Hochschulen verfügen private Hochschulen im Allgemeinen über ein sehr eingeschränktes und spezialisiertes Spektrum an Studiengängen (*Levy 1992; Stannek/Ziegele 2005, S. 23 ff.; Wells/Sadlak/Masceanu 2007, S. 513 ff.*). Ihre Überlebenswahrscheinlichkeit müsste somit größer als die der öffentlichen Hochschulen sein.² Aus Befunden der organisationsökologischen Studien lässt sich also eine Annahme ableiten, die die neoliberale Erwartung stützt, nach der private Hochschulen einen Überlebensvorteil gegenüber öffentlichen haben.

Vor dem Hintergrund dieser Annahmen und Erwartungen wird im Folgenden untersucht, welche Organisationsmerkmale das Überleben von Hochschulen beeinflussen. Es geht vor allem um die Frage, ob die Trägerschaft der Hochschulen einen Einfluss auf deren Überlebenswahrscheinlichkeit hat. Daneben wird im direkten Anschluss an die organisationsökologischen Befunde auch gefragt, ob sich mit einer spezialisierten Ausrichtung der Hochschulen Überlebensvorteile verbinden. Sind beispielsweise Hochschulen mit religiösem Träger, deren Bildungsprogramme Elemente partikularen Charakters enthalten und die somit spezielle Bildungsinteressen bedienen, überlebentüchtiger als solche mit einem ausschließlich wissenschaftlich universalistischen Charakter? Ferner wurde der Frage nachgegangen, ob weitere Variablen, etwa die Anzahl der Studierenden, der Umfang des Personals etc. einen Einfluss auf die Überlebenswahrscheinlichkeit der Hochschulen haben.

2 Methoden

In methodischer Hinsicht folgen wir dem überlebensanalytischen Ansatz der *organizational ecology*. Wir haben aus aggregierten Daten der amtlichen Hochschulstatistiken und aus Individualdaten zu den Hochschulen eines Landes Zeitreihen gebildet, die die Administration (Trägerschaft, Finanzierung etc.), die Lehre (Anzahl Studierende, Anzahl und Qualifikation der Lehrenden etc.) und die Forschung der einzelnen Hochschulen betreffen. Sie wurden mit Überlebensmodellen analysiert.

² Ein ähnliches Argument folgt auch aus den Forschungen zur Hochschulentwicklung, die die Organisationsstrukturen von Hochschulen in den Mittelpunkt stellt. „Institutional specialization“, so heißt es etwa zusammenfassend bei van Vught, „allows higher education institutions to focus their attention and energy, which helps them in producing higher levels of effectiveness“ (*Vught 1996, S. 45*).

Die Überlebensanalyse zielt auf Phänomene, in deren Mittelpunkt der Zeitablauf bis zum Eintritt eines bestimmten Ereignisses steht. Untersucht werden die Zeitdauer und deren Beziehungen zu anderen Variablen. Das Grundproblem dieses Verfahrens besteht im Modellieren von Zeit-Ereignis-Daten beziehungsweise Zeit-Untergangs-Daten. Der klassische Gegenstand dieser Analyse ist die Überlebenszeit nach einer bestimmten medizinischen Behandlung, also die Zeit bis zum Tod oder einem Rückfall (*Cheung 2000*). Zentrale Begriffe sind stets Tod und Überleben, auch wenn die zu untersuchenden Ereignisse ihrem materialen Inhalt nach nicht mit biotischen Organismen oder Populationen von Organismen im Zusammenhang stehen. Untersucht werden grundsätzlich Paneldaten, d.h. Daten über eine Anzahl von individuellen Mitgliedern eines Kollektivs zu verschiedenen Zeitpunkten. Ein wesentlicher Unterschied zwischen der Überlebensanalyse und anderen statistischen Methoden zur Analyse von Paneldaten besteht in der Akzeptanz von so genannten „censored cases“. Sie beziehen sich auf Individuen oder Organisationen, deren Lebens- bzw. Bestandsentwicklung nicht über die ganze Untersuchungsperiode hinweg verfolgt wurde oder bei denen sich das interessierende Ereignis während der Untersuchungsperiode nicht einstellte. In der Arzneimittelforschung z. B. sind diese Fälle bedeutsam, da sie auf eine erfolgreiche Behandlung hindeuten.

In der vorliegenden Untersuchung kamen vor allem Cox-Regressionen (*Cox 1972*) zum Einsatz. Diese Methode verknüpft in multiplen Regressionen den Hazard von Ereignissen mit Kovariablen. Das Modell erlaubt es, den Hazard zu schätzen. Mit dem Konzept des Hazards wird die Wahrscheinlichkeit pro Zeiteinheit bezeichnet, dass eine Organisation innerhalb eines Zeitintervalls das Zielereignis (z. B. Untergang) erfährt, wenn sie denn bis zum Beginn dieses Zeitintervalls überlebt hat. Es geht also um das Risiko pro Zeiteinheit für den Untergang zur Zeit t . Mathematisch lässt sich dieses Modell folgendermaßen beschreiben:

$$h(t, (z_1, z_2, \dots, z_m)) = h_0(t) * e^{b_1 * z_1 + \dots + b_m * z_m}$$

Die Größe $h(t, \dots)$ ist der Hazard für die Werte der Einflussvariablen (Kovariablen) der Fälle (z_1, z_2, \dots, z_m) und der Überlebenszeit t . Die Größe $h_0(t)$ ist der Baseline-Hazard und gibt den Hazard für das Eintreten des Ereignisses in t an, wenn alle Einflussvariablen gleich null sind. Das Modell kann als lineare Gleichung dargestellt werden, nur muss dann der Logarithmus betrachtet werden:

$$\log[h(t, (z_1, z_2, \dots, z_m))/h_0(t)] = b_1 * z_1 + \dots + b_m * z_m$$

Drei Länder, nämlich Chile, Deutschland und die USA wurden in die Untersuchung einbezogen. Es wurden Länder ausgewählt, deren gesellschaftliche Entwicklung und deren Hochschulen sich beträchtlich unterscheiden. Auf diese Weise sollen Hinweise auf allgemeine Trends der Entwicklung herausgearbeitet werden, die sich jenseits

nationaler Besonderheiten durchsetzen. Da die nationalen Hochschulstatistiken nicht einheitlich sind, variieren die Variablen, die in die Analyse einbezogen werden konnten, in gewissen Grenzen von Land zu Land.³ Die Unterscheidung von öffentlichen und privaten Hochschulen erfolgt entsprechend der amtlichen Zuordnungen, die den nationalen Hochschulstatistiken zugrunde liegen. Diese gehen von der Trägerschaft aus. Auch die Unterscheidung von profitorientierten und gemeinnützigen (oder: non-profit) Hochschulen im privaten Bereich folgt den rechtlichen Zuordnungen in den jeweiligen Ländern.⁴

3 Ergebnisse im Ländervergleich

3.1 USA

Im Falle der USA wurden Daten zum Überleben der *4-years-colleges* für die Zeitspanne von 1980 bis 2004 ausgewertet. Es handelt sich durchweg um Colleges, die von staatlich anerkannten Agenturen akkreditiert sind.

Die Zahl der profitorientierten privaten Colleges hat in den vergangenen Jahrzehnten beträchtlich zugenommen. Sie stieg von 67 im Jahr 1986⁵ auf 447 im Jahr 2004. Von den 574 Einrichtungen, die in der genannten Zeitspanne einmal existierten, sind 75 geschlossen worden. Die Zahl der privaten gemeinnützigen Colleges stieg von 1.409 im Jahr 1980 auf 2.154 in 1996 und sank bis 2004 auf 1.827. Insgesamt hatte es im Beobachtungszeitraum einmal 2.644 gegeben. Von ihnen wurden 482 geschlossen. Die Gesamtzahl der öffentlichen Colleges betrug 639 im Jahr 1980, blieb im Großen und Ganzen bis 1996 konstant. Danach stieg sie bei kleinen Schwankungen auf 793 im Jahr 2004. Von den insgesamt 806, die einmal existierten, sind 30 geschlossen worden.

Private und öffentliche Colleges wurden während der untersuchten Zeitspanne in beachtlichen Zahlen gegründet oder geschlossen. Aus den Zeitreihendaten zur Gründung und Schließung von Hochschulen haben wir Überlebensraten⁶ berechnet. Dabei variiert die Überlebenswahrscheinlichkeit mit der Trägerschaft. Am niedrigsten war mit 0,91 (Standardfehler = 0.006) die Überlebenswahrscheinlichkeit profitorientierter

³ Eine Liste sämtlicher Datenquellen kann per E-Mail von den Autoren über das Institut für Hochschulforschung Wittenberg bezogen werden.

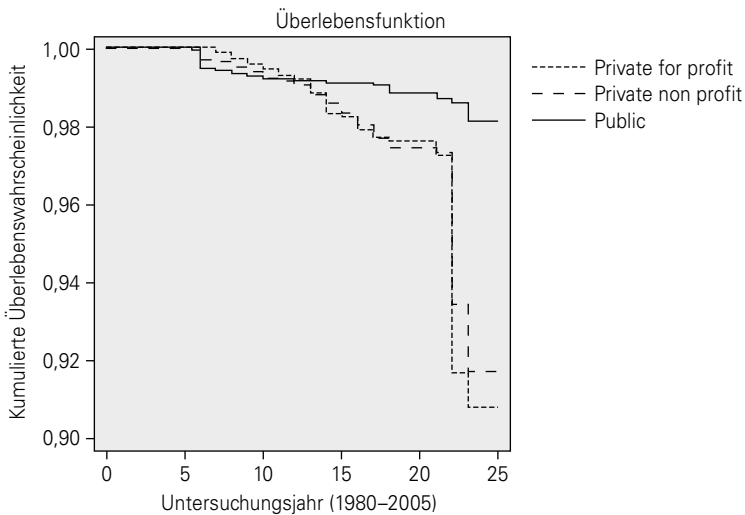
⁴ Eine ausführliche Diskussion dieser Unterscheidungen findet sich in Reisz (2003) und in Fernández/Lenhardt/Reis/Stock (2009) vor allem mit Blick auf Chile (S. 11 ff.) und auf Deutschland (S. 35 ff.) sowie mit Blick auf öffentliche und private Hochschulen in den USA (S. 80 ff.). Zu profitorientierten Hochschulen in USA vgl. Hentschke (2004).

⁵ In der Hochschulstatistik der USA werden erst ab 1986 die profitorientierten privaten Hochschulen gesondert ausgewiesen. Deshalb liegen für 1980 noch keine Daten vor.

⁶ Die Überlebensrate gibt die Wahrscheinlichkeit an, eine bestimmte Zeiteinheit zu überleben. Die Einzelwahrscheinlichkeiten werden zu einem Wert zusammengeführt.

Hochschulen. Etwas größer war mit 0,92 (Standardfehler = 0.003) die der privaten gemeinnützigen und am größten die der öffentlichen mit 0,98 (Standardfehler = 0.002). Die entsprechenden Überlebensfunktionen zeigt Abbildung 1. Es gehört demnach zu den Kennzeichen des amerikanischen Hochschulwesens, dass die Existenz öffentlicher Hochschulen sicherer ist als die der profitorientierten und die der gemeinnützigen privaten.

Abbildung 1: Kumulierte Überlebenswahrscheinlichkeit amerikanischer Hochschulen nach Trägerschaft



Datenquellen: siehe Fernández/Lenhardt/Reisz/Stock (2009: 114f.)

Zu den Bildungsinteressen, die die Hochschulen verfolgen, gehören in den USA immer noch in großem Umfang religiöse. Dabei wird die religiöse Bildung primär als Teil der bürgerlichen Bildung gepflegt und nicht als theologische Fachbildung für zukünftige Geistliche. Staatliche Hochschulen scheiden wegen der Trennung von Kirche und Staat als Ort religiöser Bildung aus. Ebenso wenig widmen sich ihr die profitorientierten Hochschulen.

Die Überlebensanalyse ergibt, dass private Colleges mit religiösen Orientierungen eine um 23 Prozent größere Überlebenschance haben als solche mit rein säkularer Orientierung (Sig.=0,009). Dieser Befund scheint die These zu bestätigen, die Befriedigung von Sonderinteressen verleihe diesen Einrichtungen günstigere Existenzbedingungen. Kontrolliert man das Alter der Colleges, verschwindet dieser Effekt religiöser Bindung jedoch. Es zeigt sich dann, dass das geringere Sterberisiko der entsprechenden Einrichtungen allein auf ihr früheres Gründungsdatum zurückgeht (Sig.=0,000). Bei jün-

geren Gründungen verbessern religiöse Bindungen die Überlebenschancen nicht. Colleges jeglicher Orientierung und jeglicher Trägerschaft haben eine deutlich größere Überlebenschance, wenn ihr Gründungsdatum länger zurückliegt. Am stärksten ist dieser Effekt bei den öffentlichen (Verbesserung um bis zu 9 Prozent; Sig.=0,000) und den privaten gemeinnützigen (Verbesserung um bis zu 6 Prozent; Sig.=0,000). Die These, partikulare Orientierungen, in diesem Fall also religiöse, begünstigten das Überleben privater Hochschulen, wird durch diese Befunde also nicht gestützt. Dass eine Hochschule mit einer längeren Tradition günstigere Überlebenschancen hat, lässt sich damit erklären, dass sie sich in den Augen der Öffentlichkeit bewährt hat. Hinzu kommt, dass eine längere Tradition gleichbedeutend ist mit einer größeren Zahl an Absolventenjahrgängen und mithin an dankbaren und zu Spenden bereiten Alumni. Da „private giving“ für die Budgets und damit für den Bestand einer Hochschule höchst bedeutsam ist, haben die älteren einen Überlebensvorteil.

Am Falle der USA kann auch untersucht werden, ob Hochschulen mit größerem wissenschaftlichem Anspruch größere Überlebenschancen haben als die übrigen. Als Indikator für den wissenschaftlichen Anspruch nehmen wir den höchsten Abschluss, den ein College vergeben darf. Der niedrigste Abschluss, den Vier-Jahres-Colleges vergeben, ist der *Bachelor*. Höher rangieren der *First Professional Degree* und der *Master* und am höchsten der Abschluss *post M.A. less than Doctorate*. Bei der Einordnung der einzelnen Hochschulen folgen wir der Kategorisierung der *Carnegie Foundation*. Sie bestimmt den Rang eines Colleges nach dem höchsten Abschluss, den die Absolventen hier jeweils erwerben können. Die Analyse ergibt: Profitorientierte Colleges haben ein größeres Risiko, schließen zu müssen, wenn sie höhere Abschlüsse vergeben. Das Schließungsrisiko steigt um 18 Prozent (Sig.=0,000). Dieser Zusammenhang ist unabhängig von der Anzahl der Studierenden. In der Sprache der Geschäftswelt formuliert bedeutet dieser Befund, dass profitorientierte Colleges am Markt vorbeiproduzieren, wenn sie eine anspruchsvollere wissenschaftliche Bildung anbieten. Nach den allgemeinen Erwartungen der Konsumenten, so lässt sich spekulieren, kommen sie dafür nicht in Frage. Deswegen ist ihr Risiko zu scheitern hoch. Dem entspricht der Befund, dass private gemeinnützige Colleges ein kleineres Sterberisiko haben, wenn sie höhere Abschlüsse vergeben. Das Sterberisiko fällt dann um 4,5 Prozent (Sig.=0,002). Nimmt man die Anzahl der Studierenden hinzu, fällt es sogar um 11 Prozent (Sig.=0,000). Jedoch wächst auch das Schließungsrisiko öffentlicher Colleges, wenn sie höhere Abschlüsse vergeben. Es nimmt um 11 Prozent zu (Sig.=0,036).

Je kleiner der Lehrkörper der Colleges desto größer ist ihr Sterberisiko. Es kann bis zu 5 Prozent ansteigen (Sig.=0,036). Das gilt nicht nur für öffentliche Colleges, sondern auch für private gemeinnützige. Hier steigt das Risiko um jeweils 4,5 Prozent (Sig.=0,000 in beiden Fällen). Eine nahe liegende Erklärung dieses Befundes lautet so: Angesichts

der fortgeschrittenen akademischen Arbeitsteilung verlangt eine anspruchsvollere wissenschaftliche Bildung spezialisierterer Professoren und Lehrveranstaltungen. Diese bereitzuhalten übersteigt aber schnell die Möglichkeiten eines kleineren Colleges, gleichviel ob es sich in privater oder öffentlicher Trägerschaft befindet. Dagegen sind sowohl die Anzahl der Studierenden wie auch die Relation zwischen der Anzahl der Lehrenden und der Lernenden bei allen Hochschulen ungeachtet ihrer Trägerschaft ohne statistische Relevanz. Vor allem die Größe des Lehrpersonals befördert das Überleben eines Colleges.

Träfe die These von der Überlegenheit privater Hochschulträgerschaft zu, müsste es einen Zusammenhang geben zwischen der Form der Trägerschaft und der Leistungsfähigkeit eines Colleges in Forschung und Lehre. Ob ein solcher Zusammenhang besteht, wurde anhand der Daten des *Center for Measuring University Performance* untersucht, das jedes Jahr einen Bericht „*Top American Research Universities*“ publiziert (Craig 2002; Lombardi/Capaldi/Abbey 2007). Das Center bezieht insgesamt 640 Hochschulen in seine Untersuchung ein. Das sind jene, die mindestens in einem Jahr zwischen 1990 und 2004 bundesstaatliche Forschungsgelder erlangt haben. Der Schwerpunkt liegt aber bei den Forschungsuniversitäten, die mindestens 20 Millionen Dollar an bundesstaatlichen Forschungsmitteln pro Jahr erlangten. Das waren im Jahre 2005 196 Universitäten, darunter 57 private und 136 öffentliche. Die zur Einstufung benutzten Indikatoren betreffen fünf Felder, nämlich Forschungsmittel (Gesamtausgaben für Forschung und eingeworbene bundesstaatliche Forschungsmittel), private Zuwendungen, die Leistungen, Mitgliedschaften und Auszeichnungen des wissenschaftlichen Personals, Förderung des wissenschaftlichen Nachwuchses (Anzahl der im akademischen Jahr erworbenen Dokortitel und Anzahl der Postdoc-Stellen) sowie das Leistungsvermögen der undergraduates. Für jeden der Indikatoren wird eine Rangliste aufgestellt, in die die Universitäten einsortiert werden. Zur ersten Gruppe der *Top American Research Universities* (1–25) werden jene Hochschulen gerechnet, die bei mindestens einem der neun Indikatoren einen der ersten 25 Plätze belegen. 2005 gelang dies 51 Universitäten. In eine zweite Gruppe werden jene aufgenommen, die mit mindestens einem Merkmal auf den Plätzen 26 bis 50 vertreten sind. Hier finden sich 37 Universitäten. Ferner gibt es Listen der Top 200, in die sämtliche 640 Hochschulen aufgenommen werden, die bundesstaatliche Forschungsmittel erlangten. Eine systematische Analyse der genannten Daten zeigt, dass die Form der Trägerschaft nicht mit der Zugehörigkeit zur Leistungsspitze korrespondiert. Unter den ersten 25 Hochschulen sind öffentliche und private zu etwa gleichen Anteilen vertreten, die Gruppe 26 bis 50 wird zu fast drei Vierteln von öffentlichen Hochschulen gebildet. Zwischen der Unterscheidung von Universitäten, die zur Spitzengruppe der 25 zählen und jenen, die nicht zur Spitze gehören einerseits sowie der Unterscheidung von öffentlicher und privater Trägerschaft andererseits gibt es insgesamt keinen Zusammen-

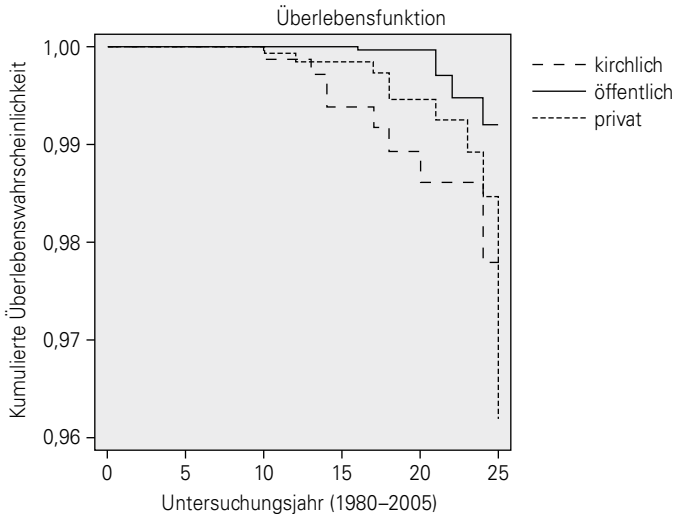
hang ($\phi=-0,18$; $\text{Sig.}=0,000$). Dieser Befund widerspricht der landläufigen Ansicht, die privaten Universitäten nähmen die Spitzenpositionen im akademischen Feld ein.

Für die USA ergibt sich damit insgesamt das folgende Bild: Privat zu sein und gar profitorientiert verschafft einem College nicht Überlebensvorteile, sondern -nachteile. Ein jüngeres Gründungsdatum ist zudem gefährlicher als ein älteres, und ein kleiner Lehrkörper riskanter als ein großer. Dagegen spielt die Größe der Studentenschaft keine Rolle und eine religiöse Bindung auch nicht. Die privaten Hochschulen sind den öffentlichen insgesamt nicht überlegen. Überlebensvorteile, die sich einer Überlegenheit der privaten gegenüber den öffentlichen Hochschulen verdanken, können also kaum entstehen.

3.2 Deutschland

Das Schicksal privater Hochschulen findet in Deutschland große Aufmerksamkeit. Über die Existenznöte der Universität Witten-Herdecke haben die Medien ausführlich berichtet, ebenso über die Schließung der Hanseuniversität in Rostock und der International University Bruchsal im Jahre 2009. Einen genaueren Blick gestatten statistische Daten, die für Überlebensanalysen geeignet sind. Sie liegen für die Jahre von 1980 bis 2006 vor. In dieser Zeitspanne existierten insgesamt 131 private Hochschulen, 49 kirchliche und 292 öffentliche. Von ihnen wurden 24 private Hochschulen geschlossen, neun kirchliche und 21 öffentliche. Aus den Zeitreihendaten zur Gründung und Schließung von Hochschulen wurden auch hier Überlebensraten berechnet. Sie betragen für die staatlichen Hochschulen 0,98, für die kirchlichen 0,96 und für private Hochschulen 0,94. Die Überlebenswahrscheinlichkeit der öffentlichen Hochschulen war signifikant höher als die der übrigen. Die Überlebensfunktionen für den deutschen Fall zeigt Abbildung 2.

Abbildung 2: Kumulierte Überlebenswahrscheinlichkeit deutscher Hochschulen nach Trägerschaft



Datenquellen: siehe Fernández/Lenhardt/Reisz/Stock (2009: 114f.)

Von den Variablen, für die Daten zur Verfügung stehen, hat das Gründungsdatum der Hochschulen einen Effekt auf die Überlebensraten, und zwar ausschließlich im Fall der kirchlichen Hochschulen. Ein frühes Gründungsdatum begünstigt die Überlebenschancen um ein Prozent, bei schwacher Signifikanz (Sig.=0,1759). Für private und für staatliche Hochschulen lässt sich ein solcher Zusammenhang jedoch nicht feststellen. Ohne Einfluss auf die Überlebenschancen ist in den Hochschulen aller Träger die Anzahl der Studierenden. Angesichts der häufig vertretenen Annahme, dass kleinere Organisationen ein größeres Risiko haben, geschlossen zu werden, ist dies ein überraschender Befund. Ohne Bedeutung für die Überlebenschance ist auch die Anzahl der Absolventen.

Hingegen erhöht der Umfang des Lehrpersonals die Überlebenschancen der kirchlichen Hochschulen um bis zu 3,5 Prozent (Sig.=0,0058). Dieser Befund gibt Auskunft über die Schließungskriterien der zuständigen Kirchenleitungen. Insbesondere nach dem Beitritt der ostdeutschen Länder gab es, etwa in Berlin, Mehrfachangebote an Theologischen Hochschulen, die dann zu Schließungen führten. Offenbar hat man sich bei der Auswahl der zu schließenden Einrichtungen vom Umfang des Lehrkörpers leiten lassen. Für private Hochschulen ist der Zusammenhang nicht signifikant.

Die anderen Variablen, die in die Untersuchung einbezogen werden konnten, haben keine signifikanten Effekte auf die Überlebenswahrscheinlichkeit. Ob die Hochschulen das Recht zur Promotion oder zur Habilitation haben, ob sie über eine größere oder

kleinere Betreuungsrate (Relation von Lehrenden und Lernenden) verfügen, ob sie einen großen Teil der Studierenden zu einem Abschluss führen oder nicht, ob der Anteil der nebenberuflich Lehrenden groß oder klein ist, ist für ihr Überleben ohne signifikante Bedeutung. Der einzige bedeutende Faktor bleibt die Trägerschaft selbst. Private Hochschulen sind schließungsgefährdeter als die kirchlichen und diese wiederum sind gefährdeter als die öffentlichen.

Ob die privaten Hochschulen privatwirtschaftlich arbeiten oder gemeinnützigen Charakters sind, ließ sich mangels statistischer Daten mit Zeitreihen nicht beschreiben. Die deutschen Hochschulgesetze schließen eine privatwirtschaftliche Organisationsform der privaten Hochschulen nicht aus (*Schmidt 2003; Erichsen 2000, S. 148 ff.*). In der untersuchten Zeitspanne von 1980 bis 2006 dürfte es allerdings kaum profitorientierte Hochschulen gegeben haben. Die erste Übernahme einer bestehenden privaten Hochschule durch einen Investor fand 2007 statt. Die Internationale Fachhochschule Bad Honnef-Bonn wurde durch den Finanzinvestor *Aucutus* übernommen. Im Jahr 2008 erwarb der amerikanische Bildungskonzern *Laureate Education* die Mehrheitsanteile an der *Unternehmer-Hochschule BiTS (Business and Information Technology School)* in Iserlohn. Mindestens sechs weitere Übernahmen oder Neugründungen durch Investoren sind für die Jahre 2007 und 2008 dokumentiert (*Savelberg 2008*). Gemeinhin galt die Hanseuniversität Rostock als eine der ersten Hochschulgründungen, die ein privater Investor in Deutschland, nämlich die Hamburger Beteiligungsgesellschaft *Educationtrend* angestrengt hat (*Gürtler/Schmid 2008*). Dies war 2008. Der gleiche Investor hatte 2007 bereits die *International University Bruchsal* erworben. Gleichwohl dürfte die anfängliche Euphorie über die Entstehung eines profitorientierten „Education Services Sektor“ (*Savelberg 2008, S. 178*) im Hochschulbereich, der Investoren anlocke, einer Ernüchterung gewichen sein. Die Hanseuniversität Rostock ist 2009 gescheitert, ebenso musste die International University Bruchsal im August desselben Jahres ihren Betrieb einstellen. Auch wenn keine systematischen hochschulstatistischen Erhebungen vorliegen, so deuten die genannten Befunde auf ein vergleichsweise hohes Untergangrisiko profitorientierter Privathochschulen in Deutschland. Es scheint derzeit höher zu sein als das der gemeinnützigen privaten Hochschulen.

Initiativen zur Gründung privater Hochschulen können auch scheitern, bevor die geplante Einrichtung ihre Tore öffnete. Bis 2001 bedurften private Hochschulen allein einer Anerkennung durch den Staat. Dann trat zur staatlichen Anerkennung das Erfordernis der institutionellen Akkreditierung. Das Verfahren wird im Auftrag des Bundes und der Länder vom Wissenschaftsrat durchgeführt. Anlass für die Einrichtung des Verfahrens war die seit den 1990er Jahren sprunghaft gestiegene Anzahl privater Hochschulen (*Bartz/Goll/Lütkemeier/Schwörer 2009*). In einigen Landeshochschulgesetzen wird die institutionelle Akkreditierung ausdrücklich als Voraussetzung staat-

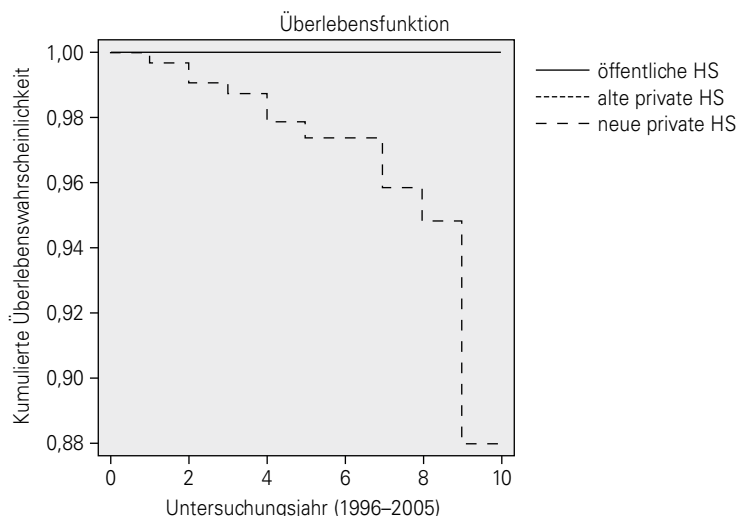
licher Zulassung genannt. Der Wissenschaftsrat hat zwischen 2001 bis 2008 im Verfahren der institutionellen Akkreditierung insgesamt 32 Stellungnahmen verabschiedet. Davon waren sechs ablehnend und 26 zustimmend. Andere Beobachter sprechen von 55 Verfahren bis Anfang 2009, wobei in 20 Fällen die Akkreditierung verweigert wurde (*Schmidt 2009, S. 15*). Der Leiter des Referats „Institutionelle Akkreditierung“ des Wissenschaftsrates nennt im Jahr 2009 insgesamt 40 abgeschlossene Verfahren. Davon endeten sechs mit einem ablehnenden Bescheid, während etwa 20 Gründungsinitiativen bereits im laufenden Verfahren aus Qualitätsgründen abgebrochen wurden (*Goll 2009, S. 6*). Die Akkreditierungen werden zeitlich befristet für eine Zeitspanne zwischen fünf bis zehn Jahren erteilt. Akkreditiert werden sowohl Konzepte für zu gründende Hochschulen als auch Hochschulen, die bereits den Studienbetrieb aufgenommen haben. Hier werden die bislang erreichten Leistungen in Lehre und Forschung bewertet. Das Akkreditierungsverfahren des Wissenschaftsrates wird in jüngster Zeit von Seiten privatwirtschaftlicher Hochschulbetreiber wie etwa der SRH Holding scharf kritisiert (*Schmidt 2009*). Das Verfahren musste sich jüngst einer Überprüfung stellen, deren Ergebnis derzeit noch offen ist.

3.3 Chile

Die chilenische Hochschulstatistik bietet Daten, die für Überlebensanalysen geeignet sind, für den Zeitraum 1996 bis 2006. Sie stehen nur für die Universitäten zur Verfügung. Insgesamt existierten im Untersuchungszeitraum 74 Universitäten, davon waren 16 staatliche Universitäten, 3 alte private und 6 alte katholische Universitäten. Diese alten nichtstaatlichen Universitäten (*Tradicional*) sind den staatlichen hochschulpolitisch gleichgestellt. Die genannten insgesamt 25 Universitäten bilden auch den *Consejo des Rectores de las Universidades Chilenas (CRUCH)*, den Universitätsrektorenrat. Unabhängig von der Trägerschaft sichert der Staat die Grundfinanzierung – bis zu zwei Dritteln des Gesamtetats – dieser Universitäten (*Lanzendorf 2005*). Daneben gab es 49 neue private Hochschulen. Diese gehören nicht dem CRUCH an.

Zwischen 1996 und 2006 wurden 14 Hochschulen geschlossen. Die Schließungszeitpunkte verteilen sich auf die Zeitspanne relativ gleichmäßig. Von diesen Schließungen waren staatliche Universitäten überhaupt nicht betroffen und auch nicht jene alten privaten und kirchlichen Universitäten, die zum CRUCH gehören. Betroffen waren ausschließlich neue private Hochschulen. Während die Universitäten des CRUCH eine Überlebenswahrscheinlichkeit von 100 Prozent haben, beträgt sie für die neuen privaten Universitäten nur 0,88. Abbildung 3 zeigt die entsprechenden Überlebensfunktionen.

Abbildung 3: Kumulierte Überlebenswahrscheinlichkeit chilenischer Hochschulen nach Trägerschaft



Datenquellen: siehe Fernández/Lenhardt/Reisz/Stock (2009: 114f.)

In Chile verliert eine Universität ihre Lizenz, wenn sie ihren Entwicklungsplan nicht einhält oder wenn sie elf Jahre nach ihrer Gründung noch nicht den Status der Autonomie erreicht hat.⁷ Vor einer Gründung müssen die Entwicklungspläne der zukünftigen Universität beim *Consejo Superior de Educación (CSE)*, dem Hochschulrat, zur Begutachtung eingereicht werden. Nach der Gründung begleitet der CSE die Universität elf Jahre und evaluiert die Studiengänge und das Management der Einrichtung. Nach dieser Prüfungszeit entscheidet der Hochschulrat, ob die Universität die angestrebte Autonomie erhalten oder aber geschlossen werden sollte. Oft kommt es vor, dass die Träger selbst entscheiden, ihre Universität aufzulösen. So haben sich von den 14 Universitäten, die geschlossen worden sind, sieben selbst aufgelöst, die anderen sieben wurden vom CSE geschlossen. Die Selbstauflösung kann durch das Interesse motiviert sein, einer Schließung durch den CSE zuvorzukommen, oder sie ergibt sich aus einem Verkauf der Universität. Hochschulen werden an- und verkauft, weil sie ungeachtet der gesetzlich vorgeschriebenen Gemeinnützigkeit häufig erhebliche Gewinne abwerfen.

Die für die Universitäten zur Verfügung stehenden Daten gestatten es, eine große Zahl an möglichen Determinanten der Überlebenschancen zu prüfen. Im Einzelnen sind dies: die Form der Trägerschaft; die Anzahl der Studiengänge; die Anzahl der Fachbereiche; die durchschnittliche Dauer der Regelstudienzeit; der Anteil der Studie-

⁷ Zur institutionellen Akkreditierung in Chile vgl. *Bernasconi (2007)*.

renden, die das erste Studienjahr erfolgreich beendeten; die Anzahl der Studienplätze, die jährlich angeboten werden; die Anzahl der pro Jahr immatrikulierten Studenten; die Anzahl der Studierenden; die Immatrikulationsgebühren; die Studiengebühren pro Jahr; die Anzahl der Professoren und der Anteil, der über einen Master oder Dokortitel verfügt; die Anzahl und der Anteil der Professoren, die mindestens eine halbe Stelle haben; die Anzahl der Studierenden, die ihr Studium schon beendet, aber noch nicht das Abschlussexamen abgelegt haben. Da alle Schließungen in den genannten Zeitraum von zehn Jahren fielen, schied das Gründungsdatum als Variable aus.

Es ist sinnvoll, den Einfluss dieser Variablen auf das Überleben sowohl aller Hochschulen zu berechnen als auch gesondert auf das der neuen privaten. Hervorzuheben sind zwei Befunde: Bezogen auf die gesamte Population hat die Anzahl der Studiengänge einen hoch signifikanten (Sig.=0.001) Einfluss auf das Überleben. Eine hohe Anzahl von Studiengängen verbessert die Überlebenschancen einer Hochschule um bis zu 6,2 Prozent. Dies gilt, wie gesagt, für die Gesamtheit aller untersuchten Universitäten, also unabhängig von deren Trägerschaft. Wird der Einfluss der Variable allein für die neuen privaten Hochschulen berechnet, also allein für die faktisch bestandsgefährdeten Hochschulen, so wachsen deren Überlebenschancen um 3,3 Prozent (Sig.=0,04). Ebenso hat die Anzahl der Fachrichtungen einen ganz erheblichen Einfluss auf das Überleben. Legt man sämtliche Hochschulen zu Grunde, so verbessert eine hohe Anzahl von Fachrichtungen die Überlebenschancen um bis zu 40,6 Prozent (Sig.=0,000). Dies ist zum Teil darauf zurückzuführen, dass jene etablierten Hochschulen, die dem CRUCH angehören und nicht schließungsgefährdet sind, sämtlich über eine hohe Anzahl von Fachrichtungen verfügen. Werden diese Hochschulen aus der Analyse ausgeklammert und nur die neuen privaten Hochschulen in die Berechnung einbezogen, so steigen deren Überlebenschancen bei einer hohen Zahl von Fachrichtungen allerdings noch immer um bis zu 36 Prozent (Sig.=0,001).

Dass die Breite des Studienangebotes in Chile einen außergewöhnlich großen Effekt auf das Überleben einer privaten Hochschule hat, erklärt sich zum einen aus deren Abhängigkeit von den Einnahmen aus Studiengebühren. Diese sind die einzige Einnahmequelle der neuen privaten Universitäten. Zum anderen wird die Einhaltung akademischer Standards bei den neuen privaten Hochschulen derzeit noch nicht durchgängig sanktioniert. Sie können sich einer Akkreditierung ihrer Studiengänge also noch entziehen. Unter diesen Voraussetzungen verbessert sich die Position einer Hochschule, wenn sie ein möglichst breites Spektrum von Studiengängen anbietet. Sie spricht damit viele Studierende an und erhöht ihre Einnahmen bei geringen zusätzlichen Kosten. Die Hochschulen agieren in der Konkurrenz um Studierende ähnlich wie Marktteilnehmer, und sie können dies, weil sie aus unterschiedlichen Gründen noch darauf verzichten können, akademische Normen einzuhalten. Für einen großen Teil der jungen Leute, die ein Studium aufnehmen wollen, sind diese Normen noch ohne

Bedeutung, da sie bislang in der Gesellschaft – im Vergleich zu den anderen untersuchten Ländern – nur schwach institutionalisiert sind. Eine Orientierung am Markt, auf die das Management chilenischer Hochschulen immer wieder hinweist, geht einher mit einem „Modernisierungsrückstand“ bezüglich der Anerkennung akademischer Normen.

Zugleich zeigen die Überlebensanalysen, dass weder die Anzahl der Studienplätze, noch die Anzahl der Studienanfänger, noch die der insgesamt an einer Hochschule Studierenden die Überlebenswahrscheinlichkeit beeinflussen. Dieses Ergebnis ist signifikant (Sig. = 0,05 und 0,02 sowie 0,02 für die genannten Variablen). Zunächst mag der Befund als paradox erscheinen, dass einerseits die Anzahl der Fachrichtungen das Überleben befördert, andererseits ein Einfluss der Studierendenanzahl signifikant aber ausgeschlossen werden kann. Ein näherer Blick auf die Geschichte der geschlossenen Universitäten zeigt aber, dass diese häufig mit einer großen Anzahl von Fachrichtungen begonnen haben. Sie konnten diese aber über die Jahre nicht aufrechterhalten und haben mehr und mehr Fachrichtungen geschlossen, ohne dass die Anzahl der Studierenden zurückging. Offenbar wurden die bereits immatrikulierten Studenten von den wenigen verbleibenden Fachrichtungen übernommen, bevor es dann zur Schließung kam. Daneben wurden Universitäten mit nur einer Fachrichtung oder ganz wenigen Fachrichtungen geschlossen, wobei die Anzahl der Studierenden vorher zumeist zurückging.

Es verbessern sich auch mit zunehmender Größe des Lehrpersonals (Sig.= 0,02) die Überlebenschancen der neuen privaten Universitäten, allerdings nur um bis zu 0,3 Prozent. Die Anzahl der Lehrenden mit Postgraduiertenabschluss verbessert die Überlebenschancen gegenüber der bloßen Größe des Lehrpersonals um bis auf das Doppelte (Sig.=0,039). Auch wenn der Einfluss relativ gering ist, so deuten diese Befunde darauf hin, dass eine Orientierung an Wissenschaftlichkeitsstandards im Interesse der Bestandserhaltung auch der neuen privaten Universitäten ist.

Neben den Universitäten gehören auch die *Institutos Profesionales* (Technische Fachschulen) und *Centros de Formación Técnica* (Technische Ausbildungszentren) zum tertiären Bereich. Allerdings liegen für die beiden zuletzt genannten Typen keine Daten vor, mit denen sich Überlebensanalysen durchführen ließen. Allein Zeitreihenangaben über die Anzahl der Einrichtungen pro Jahr sind für die Spanne von 1990 bis 2007 verfügbar. Ein Vergleich ergibt hier das Folgende: Im Jahr 1990 gab es 40 Universitäten, in der Zeitspanne von 1990 bis 2007 bewegt sich die Anzahl zwischen 35 und 45, zuletzt gibt es 36 Universitäten. Während die Anzahl der Universitäten weitgehend konstant bleibt, fällt die Anzahl der *Institutos Profesionales* kontinuierlich von 79 im Jahre 1990 auf 44 im Jahre 2007 und die der *Centros de Formación Técnica* von 161 im Jahre 1990 auf 100 im Jahre 2007. Die beiden zuletzt genannten Typen haben mit

Abstand die größten Überlebensprobleme. Diese Einrichtungen sind alle privaten Charakters mit geringem gesellschaftlichem Ansehen und einer wissenschaftlich bescheidenen Lehre. Zudem ist es ihnen im Gegensatz zu den Universitäten erlaubt, profitorientiert zu wirtschaften. So betätigt sich also auch im chilenischen Fall der Befund, dass privatwirtschaftliche Einrichtungen gefährdeter sind als gemeinnützige private und dass die öffentlichen Einrichtungen die größten Überlebenschancen haben.

4 Zusammenfassende Diskussion der Befunde

Die entscheidenden Befunde, zu denen die Untersuchung gelangt, widersprechen den eingangs angesprochenen neoliberalen Erwartungen: In allen drei Ländern ist die Überlebenswahrscheinlichkeit der privaten Hochschulen niedriger als die der öffentlichen. Wo die Daten entsprechende Berechnungen gestatten, zeigt sich, dass die Überlebenswahrscheinlichkeit der privaten Hochschulen mit erwerbswirtschaftlichem Charakter am geringsten ist. Etwas größer ist die der privaten gemeinnützigen und am größten die der öffentlichen.

Dieser Befund, so könnte man in neoliberaler Perspektive einwenden, verdanke sich der beklagenswerten Tatsache, dass der Staat Kosten-Nutzenkalküle nicht anstelle und auf Kosten des Steuerzahlers auch unrentable Hochschulen alimentiere. Private Hochschulen operierten kostengünstiger als öffentliche. Tatsächlich gibt es Hinweise darauf, dass die Kosten pro Studienplatz in privaten Hochschulen niedriger sind als in öffentlichen. Aber diese Differenz erklärt sich vor allem daraus, dass die privaten Hochschulen zumeist kaum Forschung betreiben und die Lehre auf solche Fächer beschränken, die keine großen Kosten verursachen (Levy 1992). Dieser Befund deutet auf eine irrtümliche Prämisse der neoliberalen Kritik an öffentlichen Hochschulen hin. Sie beruht auf der Annahme, dass das, was in der Hochschulentwicklung tatsächlich bedeutsam ist, identisch ist mit dem, was in den Bilanzen privater Träger als Kosten und Nutzen erscheint. Diese Gleichsetzung ist jedoch willkürlich. Denn die Leistungen und Aufwendungen des öffentlichen Hochschulwesens erschöpfen sich nicht in dem, was private Hochschulen als Kosten und als Nutzen in ihren Bilanzen berücksichtigen.

Schwieriger ist es, unsere Befunde vor dem Hintergrund der organisationsökologischen Thesen zu interpretieren, die die Überlebenschancen von spezialisierten im Vergleich zu generalistischen Organisationen betreffen. Wir konnten am chilenischen Fall zeigen, dass sich nicht mit der Spezialisierung, sondern mit der Breite des Studiengangangebotes die Überlebenswahrscheinlichkeit privater Hochschulen erhöht. Ebenso sind im deutschen Fall Hochschulen religiöser Träger, deren Bildungsprogramme Elemente partikularen Charakters enthalten und damit spezielle Bildungsinteressen bedienen, zwar überlebenstüchtiger als die privaten, aber nicht als die breit aufgestellten öffentlichen Hochschulen. Auch in den USA lassen sich die Überlebensvorteile

privater Hochschulen mit einem religiösen Träger nicht darauf zurückführen, dass diese die speziellen Interessen religiöser Gruppen ansprechen, sondern auf das frühe Gründungsdatum dieser Hochschulen. Eine Ablehnung der organisationsökologischen These folgt aus diesen Befunden aber nur dann, wenn zudem gezeigt werden kann, dass die Umwelten der Hochschulen sich tatsächlich rasch und stark wandeln. Wir hatten diese zentrale Prämisse des organisationsökologischen Arguments auch für Hochschulen unter Hinweis auf die vorherrschenden Diagnosen der Hochschulforschung unterstellt. Ob diese Prämisse tatsächlich triftig ist, muss hier aber offen bleiben. Hier müssten weitere Forschungen anschließen.

Ferner konnten wir zeigen, dass die Überlebensfähigkeit privater Hochschulen insgesamt nicht von Merkmalen abhängt, die sich mit der Qualität von Lehre und Forschung in Zusammenhang bringen lassen. Ein höherer wissenschaftlicher Anspruch begünstigt ihr Überleben nicht notwendigerweise, sondern kann es auch beeinträchtigen. Hingegen ergeben sich weitgehend unabhängig von der Trägerschaft der Hochschulen Überlebensvorteile mit zunehmendem Umfang des wissenschaftlichen Personals. Je größer der Personalumfang, desto größer scheint das soziale Kapital zu sein, das für das Überleben einer Hochschule mobilisiert werden kann, wie auch immer sich dies im konkreten Fall darstellen mag.

Literatur

Bartz, Olaf; Goll, Dietmar; Lütkemeier, Elke; Schwörer, Beatrix (2009): Institutionelle Akkreditierung von Hochschulen in nichtstaatlicher Trägerschaft in Deutschland durch den Wissenschaftsrat. In: Winfried Benz / Jürgen Kohler / Klaus Landfried (Hrsg.): Handbuch Qualität in Studium und Lehre: Evaluation nutzen – Akkreditierung sichern – Profil schärfen! [Teil] F. Akkreditierung und weitere Formen der Qualitätszertifizierung: Grundlagen und Elementarzwecke der Akkreditierung. Berlin: Raabe. F. 1.9. 1–14

Bernasconi, Andrés (2007): Chile: Accreditation versus Proliferation. In: International Higher Education 47. http://www.bc.edu/bc_org/avp/soe/cihe/newsletter/Number47/p18_Bernasconi.htm (Zugriff 9. September 2010)

Cheung, Yin Bun (2000): Martial status and mortality in British woman. In: International Journal of Epidemiology 29, S. 93–99

Cox, David R. (1972): Regression models and life tables (with discussion). Journal of the Royal Statistical Society 34, S. 187–220

Craig, Diane D. (2002): The Center Top American Research Universities: An Overview. Gainesville, FL: The Center

Fernández Darras, Enrique; Lenhardt, Gero; Reisz, Robert D.; Stock, Manfred (2009): Private Hochschulen in Chile, Deutschland, Rumänien und den USA. HoF-Arbeitsbericht 3/09. Institut für Hochschulforschung Wittenberg

Erichsen, Hans-Uwe (2000): Gutachten. Zulässigkeit, Grenzen und Folgen der Hochschulprivatisierung. Centrum für Hochschulentwicklung. Gütersloh

Goll, Dietmar (2009): Stachel im Fleisch? In: Frankfurter Allgemeine Zeitung vom 02.04.2009, S. 6

Gürtler, Lena; Schmidt, Marion (2008): Kein Student, nirgends. In: Financial Times Deutschland vom 13.08.2008, S. 8

Hannan, Michael T.; Freeman, John (1989): Organizational Ecology. Cambridge; Mass

Hannan, Michael T.; Carrol, Glenn R. (1992): Dynamics of organizational populations : density, legitimation, and competition. New York / Toronto.

Hentschke, Guilbert C. (2004): U.S. For-Profit Postsecondary Institutions: Departure or Extension? In: International Higher Education 35. http://www.bc.edu/bc_org/avp/soe/cihe/newsletter/News35/text009.htm (Zugriff: 4. Februar 2010).

Lanzendorf; Ute (2005): Rahmenbedingungen für deutsche Hochschulaktivitäten. Länderanalyse Chile. Wissenschaftliches Zentrum für Berufs- und Hochschulforschung. Universität Kassel. Kassel

Levy, Daniel C. (1992): Private Institutions of Higher Education. In: B. R. Clark & G. Neave (Eds.): The Encyclopedia of Higher Education. Oxford, S. 1183–1195

Lombardi, John V. / Capaldi, Elizabeth D. / Abbey, Craig W. (2007): The Top American Research Universities. 2007 Annual Report. The Center for Measuring University Performance at Arizona State University

Reisz, Robert D. (2003): Public Policy for Private Higher Education in Central and Eastern Europe, (Arbeitsberichte 2'03). Hrsg. von HoF Wittenberg - Institut für Hochschulforschung an der Martin-Luther-Universität Halle-Wittenberg. Wittenberg

Riddle, Phyllis (1996): The university and political Authority: Historical Trends and Contemporary Possibilities. In: Research in Sociology of Education and Socialization 11, S. 43–62.

Savelberg, Albert H. (2008): M&A im Education Services Sektor: Zukunftsbranche lockt Investoren. In: M&A Review. Heft 4, S. 178–185

Schmidt, Karsten (2003): Hochschulen in Rechtsformen des privaten Rechts. In: Jörn Axel Kämmerer; Peter Rawert (Hg.): Hochschulstandort Deutschland. Rechtlicher Rahmen – Politische Herausforderungen. Köln, S. 105–118

Schmidt, Marion (2009): Ärger mit dem Hochschul-TÜV. In: Financial Times Deutschland vom 27.02.2009. S. 15

Skilbeck, Malcolm (1997): Higher Education in a Changing Environment. Regional, National and Trans-National Issues. In: Tertiary Education and Management, Vol. 3, No.2, S. 101–111

Stannek, Antje; Ziegele, Frank (2005): Private Higher Education in Europe. A National Report on Germany. CHE Arbeitspapier 71. Gütersloh

Vught, F. van (1996): Isomorphism in Higher Education? In: Lynn Meek; Leo Goedegebuure; Osmo Kivinen; Risto L. Rinne (Eds.), *The Mockers and the Mocked: Comparative Perspectives on Differentiation, Convergence and Diversity in Higher Education*. Oxford, S. 42–58

Wächter, Bernd (2004): Higher Education in a Changing Environment. Internationalisation of Higher Education Policy in Europe. Bonn

Wells, Peter James; Sadlak, Jan; Vlăsceanu, Lazăr (2007): The Rising Role and Relevance of Private Higher Education in Europe. UNESCO-CEPES. Bucharest

Anschriften der Verfasser:

Dr. Gero Lenhardt

Freier Mitarbeiter am Institut für Hochschulforschung Wittenberg an der Martin-Luther-Universität Halle-Wittenberg

Collegienstraße 62

06886 Wittenberg

E-Mail: gero.lenhardt@hof.uni-halle.de

Prof. Dr. Robert Reisz

Professor für Sozialstatistik an der West Universität Timisoara

und wissenschaftlicher Mitarbeiter am Institut für Hochschulforschung Wittenberg an der Martin-Luther-Universität Halle-Wittenberg

Collegienstraße 62

06886 Wittenberg

E-Mail: reisz@info.uvt.ro

PD Dr. Manfred Stock

Vertretungsprofessor für Bildungssoziologie an der Martin-Luther-Universität Halle-Wittenberg und wissenschaftlicher Mitarbeiter am Institut für Hochschulforschung Wittenberg an der Martin-Luther-Universität Halle-Wittenberg

Collegienstraße 62

06886 Wittenberg

E-Mail: manfred.stock@hof.uni-halle.de

Der Journal Impact Factor: Aussagekraft, Grenzen und Alternativen in der Forschungsevaluation

Werner Marx, Lutz Bornmann

Der Journal Impact Factor ist neben dem h index zum wohl bekanntesten bibliometrischen Indikator zur Bewertung von Forschungsleistung geworden. Ursprünglich zur Einstufung von Zeitschriften als Ganzes innerhalb enger Fachkategorien erfunden, wird er inzwischen vielfach zur Bewertung der Publikationen von Wissenschaftlern und Forschungsinstituten herangezogen - was jedoch nicht mit bibliometrischer Bewertung an sich gleichgesetzt werden darf. Dieser Beitrag diskutiert zunächst die Grenzen der Anwendung von Journal Impact Factors im Hinblick auf den ursprünglichen Zweck. Es wird aufgezeigt, warum diese selbst im Bereich der Bewertung von Zeitschriften problematisch sein können. Sodann wird dargelegt, warum Journal Impact Factors für die Bewertung der Publikationen von einzelnen Wissenschaftlern oder ganzer Forschungseinrichtungen ungeeignet sind. Schließlich werden alternative und aussagekräftigere Metriken für die quantitative (d.h. bibliometrische) Forschungsevaluation aufgezeigt. Abschließend wird grundsätzlich auf die Aussagekraft und die Grenzen der Bibliometrie in der Forschungsevaluation eingegangen.

1 Einleitung

Für Verlage wissenschaftlicher Zeitschriften kommt alljährlich die Stunde der Wahrheit: Seit 1975 werden in den Journal Citation Reports (JCR) von Thomson Reuters (dem früheren Institute for Scientific Information, ISI) die Zeitschriften nach ihrem Journal Impact Factor (JIF)¹ in Ranglisten eingestuft, welche jährlich aktualisiert werden (*Thomson Reuters, 2012b*). Der Journal Impact Factor einer Zeitschrift entspricht der mittleren Anzahl der Zitierungen innerhalb eines Jahres von Veröffentlichungen dieser Zeitschrift rund zwei Jahre nach deren Publikation. Journal Impact Factors wurden ursprünglich zur Bewertung von Zeitschriften innerhalb enger Fachkategorien erfunden und zur Zusammenstellung einer Auswahl der wichtigsten Zeitschriften als Grundlage des Science Citation Index (SCI) gebraucht. Inzwischen ist aus dem zunächst wenig bekannten Hilfsmittel für Datenbankhersteller und Bibliothekare einer der bekanntesten bibliometrischen Indikatoren überhaupt geworden. Dem entsprechend ist auch die Anzahl von Publikationen über diesen Indikator drastisch angestiegen.

¹ Die Autoren verwenden bewusst das Kürzel JIF für Journal Impact Factor anstatt des häufig benutzten Kürzels IF für Impact Factor, um damit explizit den Bezug auf Zeitschriften auszudrücken und den JIF von anderen Impact Factors (z.B. personenbezogenen) zu unterscheiden.

Der Journal Impact Factor wird inzwischen vielfach auch zur Bewertung einzelner Publikationen sowie zur Evaluierung von Forschern und ganzer Forschungsinstitute herangezogen. Der verkürzte Begriff „Impact Factor“ (IF) wird deshalb oft mit zitierungsbasierter Wirkung überhaupt gleichgesetzt, was immer wieder zu erheblichen Missverständnissen und Verwirrungen führt. Die Heranziehung von Journal Impact Factors zur bibliometrischen Evaluierung von einzelnen Arbeiten bzw. Forschern bewegt sich allerdings auf dünnem Eis: Wegen der schiefen Verteilung der Zitierungen auf die Arbeiten einer Zeitschrift (wenige Arbeiten werden häufig zitiert und viele selten oder überhaupt nicht) sind die durchschnittlichen Zitationszahlen in Form der Journal Impact Factors für die Masse der Arbeiten einer Zeitschrift untypisch. Der Journal Impact Factor sagt deshalb nur wenig über den tatsächlichen Citation Impact einer einzelnen darin erschienenen Arbeit aus.

In diesem Beitrag geben wir einen Überblick über häufig genannte kritische Einwände zum Journal Impact Factor und diskutieren die Grenzen seiner Anwendung im Hinblick auf den ursprünglichen Zweck. Wir werden darlegen, dass die Heranziehung des Journal Impact Factors nicht mit bibliometrischer Bewertung an sich gleichgesetzt werden darf, dass der Journal Impact Factor selbst im dafür vorgesehenen Bereich der Bewertung von Zeitschriften problematisch sein kann und insbesondere dass dieser Indikator für die Bewertung der Publikationen von einzelnen Wissenschaftlern oder ganzer Forschungseinrichtungen ungeeignet ist. An Hand alternativer und aussagekräftigerer Metriken für die quantitative (d. h. bibliometrische) Forschungsevaluation zeigen wir auf, dass der Journal Impact Factor für personen- und institutionenbezogene Bewertungen inzwischen überflüssig geworden ist. Am Beispiel des Journal Impact Factors behandeln wir die Aussagekraft und die Grenzen der Bibliometrie in der Forschungsevaluation insgesamt und diskutieren mögliche Verzerrungen, die bei der Interpretation bibliometrischer Daten berücksichtigt werden sollten.

In den letzten Jahren sind bereits ein Vielzahl von Überblicksarbeiten erschienen, die sich mit den Vor- und Nachteilen des Journal Impact Factors sowie der Forschung zu diesem Indikator beschäftigt haben (*Adler, Ewing, Taylor, & Hall, 2009; Amin & Mabe, 2003; Bornmann, Marx, Gasparyan, & Kitas, in press; Garfield, 2006; Glänzel & Moed, 2002; Kaltenborn & Kuhn, 2003; Seglen, 1997; Thomson Reuters, 2012a; Vanclay, in press*). In diesem Beitrag werden zunächst die Kernaussagen dieser Arbeiten zusammenfassend dargestellt und diskutiert.

2 Datengrundlage

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

den Geistes- und Sozialwissenschaften. Analog zu den unter dem Web of Science (WoS) bereitgestellten Zitierindexen von Thomson Reuters werden auch in den Journal Citation Reports nicht alle gegenwärtig erscheinenden Zeitschriften einbezogen, sondern lediglich eine strenge Auswahl, schätzungsweise nur 10 Prozent aller weltweit verfügbaren wissenschaftlichen Zeitschriften (die sogenannten Source Journals). Das ist jedoch für viele Belange ausreichend, da gezeigt werden kann, dass die Wissenschaft von nur relativ wenigen Zeitschriften bestimmt wird (*Bradford, 1934*). Auswahlkriterien für die Journal Citation Reports sind die internationale Reputation der Editoren, die internationale Abdeckung des jeweiligen Fachgebietes, das regelmäßige Erscheinen, ein funktionierendes Qualitätssicherungsverfahren und weitere, zum Teil mehr subjektive Kriterien. Eine in den Journal Citation Reports nicht berücksichtigte Zeitschrift muss deshalb keine schlechte Zeitschrift sein.

3 Definition und Bestimmungsverfahren

Zur Bestimmung der Journal Impact Factors werden die Publikationen einer Zeitschrift jeweils aus einem Zeitfenster von zwei Jahren herangezogen und deren Zitierungen innerhalb des darauf folgenden Jahres ermittelt. Die Anzahl der Zitierungen wird dann durch die Anzahl der zitierbaren Publikationen (das sind Articles, Reviews und Short Communications) aus den beiden vorangegangenen Jahren dividiert. Zur Bestimmung beispielsweise des JIF-2010 wird zunächst die Anzahl der (zitierbaren) Publikationen aus den Jahren 2008 und 2009 festgestellt. Dann wird untersucht, wie oft der jeweilige Zeitschriftenname (Zeitschriftentitel) in den diversen zitierten Schreibweisen in Kombination mit den Publikationsjahren 2008 und 2009 in den Referenzen der Publikationen des Jahres 2010 erschienen ist. Schließlich wird der Quotient aus den gezählten Publikationen und deren Zitierungen gebildet. Etwas vereinfacht ausgedrückt kann man sagen, dass der Journal Impact Factor einer Zeitschrift die mittlere Anzahl der Zitierungen innerhalb eines Jahres von Veröffentlichungen dieser Zeitschrift circa zwei Jahre nach deren Publikation angibt. Die Journal Citation Reports bieten inzwischen zusätzlich auch Journal Impact Factors auf der Basis eines fünfjährigen (statt des klassischen zweijährigen) Publikationszeitraums an. So basiert z. B. der fünfjährige JIF-2010 auf den Publikationen aus den Jahren 2005 bis 2009 und deren Zitierungen im Jahre 2010.

Da für die Berechnung des Journal Impact Factors nicht die kompletten Referenzen der Arbeiten (also z. B. die Autorennamen, Publikationsjahre, Zeitschriftennamen und Seitenangaben) herangezogen werden, sondern nur die Häufigkeit des Vorkommens der Zeitschriftennamen in den Referenzen festgestellt wird, spricht man auch von Journal Title Matching. Dieses Verfahren hat den Vorteil, dass auch die falsch geschriebenen Zitierungen berücksichtigt werden, da die Autorennamen und die numerischen Angaben der Referenzen unberücksichtigt bleiben. Die häufigsten Fehler

passieren den zitierenden Autoren nicht bei den Zeitschriftentiteln, sondern bei den Autorennamen und vor allem bei den Seitenangaben. An Hand der Zitat-Varianten von viel-zitierten Arbeiten kann man empirisch zeigen, dass im Mittel rund 5 Prozent aller im SCI abgespeicherten Referenzen bezüglich der darin enthaltenen Zahlenangaben (wie z. B. der Seitenangabe) falsch geschrieben sind.

4 Verzerrungsmöglichkeiten

4.1 Fachunterschiede

Der fachübergreifende Charakter der Zitierindexe (z. B. des Web of Science) bedingt ein fachübergreifendes Klassifikationsschema, das sich nicht auf die Inhalte einzelner Publikationen, sondern auf die fachliche Einstufung der kompletten Zeitschriften durch den Datenbankhersteller (z. B. Thomson Reuters) bezieht. Die Zeitschriften des Journal Citation Reports sind nach dem gleichen Schema wie das Web of Science klassifiziert und in 174 Fachkategorien (Subject Categories – z. B. Chemistry, Analytical) eingeteilt, die zu 22 Disziplinen (Subject Areas – z. B. Chemistry) zusammengefasst werden können. Die verschiedenen Disziplinen und deren Fachkategorien zeichnen sich durch ganz unterschiedliche Publikations- und Zitiergewohnheiten bzw. -kulturen aus: Wenn man z. B. die Zitationsraten (mittlere Anzahl der Zitierungen pro Publikation) in den verschiedenen Disziplinen vergleicht, findet man einen Unterschied von fast einem Faktor zehn: Die Mathematik und die Informatik weisen die niedrigste Zitationsrate auf, die Molekularbiologie und Genetik die höchste und z. B. die Chemie sowie die Physik liegen in etwa dazwischen (*Thomson Reuters, 2012a*).

Eine wichtige Ursache für die verschiedenen Zitationsraten ist die unterschiedliche mittlere Anzahl der Referenzen im Anhang der Arbeiten aus den verschiedenen Disziplinen. Die Unterschiede in der Anzahl der Referenzen resultieren sowohl aus der Anzahl der potenziell zitierbaren Publikationen als auch aus den unterschiedlichen Zitiergewohnheiten in den verschiedenen Fächern. Die Fachunterschiede sind so groß, dass z. B. eine Zeitschrift an der Spitze des Ranking in der einen Fachkategorie in einer anderen Kategorie mit dem gleichen Journal Impact Factor im unteren Drittel des Ranking erscheinen kann. Eine weitere Ursache für die Unterschiede in den Zitationsraten besteht darin, dass die Zitiergewohnheiten innerhalb der einzelnen Fachgebiete zeitlichen Trends unterworfen sein können: Einerseits kann man den Trend beobachten, dass – bei gleicher Größe des Zitierfensters – für jüngere Arbeiten mehr Zitierungen als für ältere Arbeiten zu erwarten sind; andererseits gibt es über die Zeit Verschiebungen bei den Hot Topics in den Disziplinen. Weiterhin unterscheiden sich Fachgebiete hinsichtlich der Geschwindigkeit des Publizierens. Publikationen dynamischer Gebiete, wie einige Biowissenschaften, erhalten einen verhältnismäßig hohen Anteil ihrer Zitierungen bereits in den ersten Jahren nach der Publikation und sind durch die JIF-Formel begünstigt. Andere Gebiete, wie die Sozialwissenschaften, entfalten ihre

Wirkung erst später. Die genannten Fach- und Zeitunterschiede bei den Zitationsraten führen dazu, dass grundsätzlich nur solche Zeitschriften miteinander verglichen werden können, die zur gleichen Fachkategorie und zum gleichen JCR-Jahrgang gehören.

4.2 Unzureichende Abdeckung

Die Geistes- und Sozialwissenschaften leiden darunter, dass die Abdeckung der Publikationen durch die JCR-Zeitschriften unzureichend ist (Archambault & Gagné, 2004; Archambault, Vignola-Gagne, Cote, Lariviere, & Gingrasb, 2006). Die Wissenschaftler dieser Disziplinen publizieren oft nur einen kleineren Anteil ihrer Arbeiten in den erfassten Kernzeitschriften. Viele ihrer Publikationen erscheinen in Büchern oder in Zeitschriften, die wegen ihrer engen fachlichen Ausrichtung, ihrer lediglich nationalen Bedeutung oder aus anderen Gründen im Web of Science nicht berücksichtigt werden. In geringerem Ausmaß gilt die unzureichende Abdeckung der Publikationen auch für die technikorientierten Disziplinen wie z. B. die Computer- und Materialwissenschaften. Die Begrenzung auf einen Satz von Kernzeitschriften impliziert nicht nur eine unvollständige Erfassung der zu bewertenden Publikationen (die Zitierungen der nicht erfassten Arbeiten fallen komplett weg), sondern schneidet auch einen Teil der zitierenden Arbeiten ab: Es fehlen alle Zitierungen von zitierenden Publikationen, deren Zeitschriften nicht berücksichtigt werden.

4.3 Zeitschriftentyp und -umfang

In den Journal Citation Reports wird nicht zwischen Zeitschriftentypen wie Letter-Zeitschriften (z. B. *Physical Review Letters*) und Review-Zeitschriften (z. B. *Reviews of Modern Physics*) unterschieden. Letters und Reviews zeigen jedoch ein stark unterschiedliches Verhalten bei Zitierungen, sowohl was die mittlere Anzahl der Zitierungen als auch deren zeitlichen Verlauf betrifft: Reviews verbuchen Zitierungen, die eigentlich den Originalarbeiten zukommen sollten und werden deshalb im Verhältnis zu Zeitschriften mit klassischen Artikeln häufiger und auch länger anhaltend zitiert. Letters werden dagegen deutlich schneller zitiert, da sie eher vorläufigen Charakter haben (einem Letter sollte eigentlich ein klassischer Artikel folgen). Da die JIF-Formel Zitierungen von Publikationen nur bis zu einem Alter von maximal drei Jahren berücksichtigt, wird das Maximum der Resonanz von Letters in der Regel zwar noch erfasst, das von klassischen Artikeln und vor allem von Reviews häufig jedoch nur unzureichend. Zeitschriften, die auf jeweils andere Dokumenttypen fokussieren, sind auf Grund der genannten Unterschiede kaum miteinander vergleichbar.

Als Mittelwerte sind auch die Journal Impact Factors statistischen Schwankungen unterworfen. Neben den (in der JIF-Formel festgelegten) Zeitfenstern der erfassten Publikationen und ihrer Zitierungen hat der Umfang einer Zeitschrift (gemessen nach

der Anzahl der Publikationen pro Jahr) einen wesentlichen Einfluss auf ihre jährliche JIF-Schwankung. Zeitschriften mit weniger als 35 Artikeln können von Jahr zu Jahr um bis zu 40 Prozent beim Journal Impact Factor schwanken, solche mit mehr als 150 Artikeln schwanken dagegen beim Journal Impact Factor pro Jahr nur um rund 15 Prozent (Amin & Mabe, 2003). Damit eine durch veränderte Resonanz (bzw. veränderte Nachfrage bei den Wissenschaftlern) bedingte Schwankung signifikant ist, muss sie je nach Umfang einer Zeitschrift unterschiedlich ausfallen: Der Anstieg des Journal Impact Factors von beispielsweise 1 auf 3 im Falle einer wenig umfangreichen Zeitschrift ist noch kein signifikanter Anstieg ihrer Wirkung; im Falle einer umfangreichen Zeitschrift hingegen durchaus. In diesem Zusammenhang möchten wir auch darauf hinweisen, dass es eigentlich irreführend ist, in den Journal Citation Reports die Journal Impact Factors auf drei Stellen hinter dem Komma anzugeben. Bei Zeitschriften mit Journal Impact Factors, die sich auf einer Stelle hinter dem Komma unterscheiden, kann man kaum von einem signifikanten Unterschied beim Citation Impact ausgehen. Dies signalisiert eine höhere Genauigkeit, als sie vom Verfahren her gegeben ist. Der Grund ist die (vermeintliche) Notwendigkeit, die Zeitschriften nach ihrem Journal Impact Factor zu ranken und dabei auf einzelnen Rängen Mehrfachbesetzungen zu vermeiden.

4.4 Zähler-Nenner-Problem und Zeitfenster

Ein weiteres Problem der JIF-Formel wird als Zähler-Nenner-Problem bezeichnet: Bedingt durch das Verfahren des Journal Title Matching werden die Zitierungen der Publikationen aller Dokumenttypen (Zähler) mit der Anzahl zitierbarer Publikationen (Citable Items) (Nenner) in Relation gesetzt. Als zitierbare Publikationen gelten: Articles, Reviews und Short Communications, nicht hingegen Editorials, Letters, News und Meeting Abstracts. Die Zitierungen der nicht-zitierbaren Publikationen werden jedoch durch die Anwendung der Methode des Journal Title Matching immer mitgezählt. Bei den meisten Zeitschriften ist dies ein vernachlässigbares Problem, doch in einigen Fällen kann der Journal Impact Factor dadurch um bis zu 40 Prozent „künstlich“ ansteigen (Amin & Mabe, 2003). Vor allem einige fachübergreifende Zeitschriften, allen voran *Nature* und *Science*, werden dadurch aufgewertet: Die Zitierungen der in diesen Zeitschriften besonders zahlreichen nicht zitierbaren Publikationen tragen hier in erheblichem Umfang zum Zähler bei, werden aber im Nenner nicht mitgezählt.

Die klassische JIF-Formel berücksichtigt ein relativ kurzes Zitierungs-Zitierfenster und lässt nur eine mittlere Publikationslaufzeit von zwei Jahren zu. Diese Festlegung resultiert vermutlich aus der Absicht, eine möglichst gegenwartsnahe Aussage über die Performance einer Zeitschrift zu treffen. Damit wird aber nur ein kleiner Ausschnitt der Zeitentwicklung der Zitierungen erfasst, der über die nach vielen Jahren (oder gar nach Jahrzehnten) gemessene Wirkung wenig aussagt. Die Zitierungen der meisten Arbeiten

erreichen nach zwei bis drei Jahren ein Maximum und fallen dann mehr oder weniger schnell ab. Das gilt aber nur für die naturwissenschaftlichen Disziplinen – in manchen Disziplinen der Geisteswissenschaften wird das Maximum der Wirkung beispielsweise erst nach einem Jahrzehnt erreicht. Die Heranziehung des Journal Impact Factors auf der Basis fünfjähriger Publikationsfenster bietet diesbezüglich Vorteile, da es die jährlichen Schwankungen dämpft. Allerdings ist dies mit einem Verlust an Aktualität verbunden: Der älteste berücksichtigte Jahrgang mit der längsten Laufzeit der Zitierungen und entsprechend größerem Gewicht liegt immerhin sechs Jahre zurück.

4.5 Weitere Verzerrungen und Fehlerquellen

Die in den Journal Citation Reports aufgenommenen Zeitschriften spiegeln die Dominanz und Qualität US-amerikanischer Forschung in vielen Fachgebieten wider, aber auch die Vorliebe der Wissenschaftler (nicht nur amerikanischer) für englischsprachige Zeitschriften. Man hat verschiedentlich vermutet, dass (1) der insgesamt große Anteil amerikanischer Publikationen in der Forschung, speziell in dynamischen Forschungsgebieten (wie einigen Biowissenschaften), und (2) die Zitierung vorzugsweise von Kollegen des eigenen Landes den Journal Impact Factor amerikanischer Zeitschriften ansteigen lässt. Eindeutige empirische Belege für diese Vermutungen fehlen allerdings. Hier lässt sich ein potentieller Bias nur schwer von einem Qualitätseffekt der Beiträge trennen. Die Bevorzugung inländischer Zeitschriften ist auch keine Form des Chauvinismus, sondern hängt mit der Sprache der Veröffentlichungen bzw. der Verfügbarkeit der Zeitschriften zusammen: Die Amerikaner finden in den amerikanischen Zeitschriften ausreichend zitierbare Publikationen.

Die Methode des Journal Title Matching hat zwar ihre Vorteile (siehe oben); sie ist aber auch eine nicht vernachlässigbare Quelle von Fehlern bei der Erstellung der JCR-Daten. Autoren zitieren nicht immer den korrekten Zeitschriftennamen. Die verschiedenen Varianten, die von den Autoren in den Literaturverzeichnissen ihrer Arbeiten aufgeführt sind, werden ohne Nachbearbeitung in die Zitiereindeixe übernommen. Dadurch können die Grenzen zwischen zwei Zeitschriften mit ähnlichen Namen fließend sein (z. B. *Chemical Physics* und *Journal of Chemical Physics*). Bei bestimmten Namensvarianten kann Unklarheit darüber bestehen, welcher Zeitschrift die Zitierungen zugeordnet werden sollen. Grundsätzlich kann man feststellen, dass die Wahrscheinlichkeit von Fehlern bei der Berechnung des Journal Impact Factors mit der Komplexität der Geschichte und des Namens einer Zeitschrift ansteigt.

5 *Nature* und *Science*

Fachübergreifende Zeitschriften wie *Nature* und *Science* sind in die Fachkategorie „Multidisciplinary“ eingeordnet. Die darin erscheinenden Arbeiten stammen aus ganz

unterschiedlichen Fachdisziplinen mit stark voneinander abweichenden Zitierkulturen. Da *Nature* und *Science* quantitativ von den Biowissenschaften dominiert werden und diese Disziplinen auch die höchsten Zitationsraten aufweisen, ist auch der Journal Impact Factor dieser Zeitschriften entsprechend hoch.

Der publizierende Forscher sollte sich darum bemühen, über eine optimale Auswahl der Zeitschrift seine Fachkollegen als Zielgruppe (Leser) zu erreichen. Eine thematisch passende Zeitschrift mit einem vergleichsweise niedrigeren Journal Impact Factor kann dazu aber besser geeignet sein als eine Zeitschrift mit einem höheren Journal Impact Factor, in der das Erscheinen bestimmter Forschungsergebnisse kaum erwartet wird. Die renommierten multidisziplinären Zeitschriften wie *Nature* oder *Science* sind wegen ihrer hohen Journal Impact Factors sehr attraktive Publikationsmedien. Doch die hohen Journal Impact Factors dieser Zeitschriften resultieren nicht zuletzt aus dem Übergewicht von Artikeln aus den biologisch-medizinischen Fachbereichen und den hohen mittleren Zitationsraten der entsprechenden Publikationen. Die Publikationen aus anderen Fachgebieten entfalten eine deutlich geringere Wirkung, als es die Journal Impact Factors suggerieren. Man kann also nicht automatisch hohe Zitationszahlen erwarten, wenn man in *Science* oder *Nature* publiziert. Der Zweck dieser Zeitschriften besteht darin, grundlegende Forschungsergebnisse mit fachübergreifender Bedeutung einem breiten Leserkreis vorzustellen. Die hohe Verbreitung der Zeitschriften begünstigt zweifellos auch die Sichtbarkeit der spezielleren Arbeiten und es ist deshalb verständlich, dass viele Autoren diesen Vorteil ausnutzen möchten. Der eigentliche Sinn und Zweck multidisziplinärer Zeitschriften wird dadurch jedoch unterlaufen.

6 Missbrauch durch Zeitschriftenredaktionen

So alt wie die Geschichte des Weinbaus ist auch die Geschichte der Weinpanscherei. Das trifft leider auch auf die Bibliometrie zu: Die Wichtigkeit Publikationen-basierter quantitativer Indikatoren für finanzielle Mittel und Reputation begünstigt ihren Missbrauch sowohl durch publizierende Forscher als auch durch Zeitschriftenredaktionen (Reedijk, 2011; Wilhite & Fong, 2012). Die wichtigsten bisher beobachteten und diskutierten Methoden des Missbrauchs von Redaktionen sind:

- Die Einwirkung auf Autoren, möglichst Zeitschriften-Eigenzitate vorzunehmen.
- Die Verschiebung der regulären Publikation vielversprechender Arbeiten auf den Anfang des folgenden Jahres, damit das JIF-Zitierfenster die Anlaufphase nach der elektronischen Publikation ausklammert und in die Boom-Phase der Zitierungen fällt.

- Die gezielte Aufnahme von Zitaten in Editorials, die auf Publikationen in der eigenen Zeitschrift (oder in anderen Zeitschriften des gleichen Verlages) verweisen, welche innerhalb der vorangegangenen zwei Jahre erschienen sind.
- Die Aufnahme von Geister-Zitaten: Bei Geister-Zitaten spielt nicht der kognitive Einfluss der zitierten auf die zitierende, sondern ausschließlich die Zitierung einer Publikation aus der eigenen Zeitschrift eine Rolle.
- Die gezielte Veränderung von Dokumenttypen bei bestimmten Arbeiten (z. B. die Einstufung bestimmter Articles in die Kategorie Editorial), um die Anzahl der Citable Items und damit den Nenner in der JIF-Formel zu verkleinern.
- Die verstärkte Publikation von Reviews (Reviews werden in der Regel häufiger zitiert als Publikationen anderen Dokumententyps).

Diese Reaktionen auf bibliometrische Indikatoren wird man sicherlich niemals vollständig verhindern können (*Bornmann, 2010*). Man sollte sie jedoch berücksichtigen, wenn man bibliometrische Auswertungen vornimmt.

7 Einsatz in der Forschungsevaluation

7.1 Fragwürdige Bewertung von Forschern

Während der Journal Impact Factor früher ausschließlich für die Bewertung von Zeitschriften verwendet wurde, handelt es sich inzwischen um den wohl bekanntesten bibliometrischen Indikator, der vielfach auch zur Bewertung einzelner Publikationen sowie zur Evaluierung von Forschern oder ganzer Forschungsinstitute herangezogen wird. Der verkürzte Begriff „Impact Factor“ wird oft sogar mit zitierungsbasierter Wirkung überhaupt gleichgesetzt. Der Einsatz von Journal Impact Factors zur Bewertung einzelner Arbeiten bzw. den Arbeiten eines Forschers ist allerdings fragwürdig. Allenfalls in der Frühzeit der Bibliometrie, als andere Indikatoren kaum zur Verfügung standen, konnte man dafür noch ein gewisses Verständnis haben.

Die Tatsache, dass eine bestimmte Arbeit in einer Zeitschrift mit einem hohen oder niedrigen Journal Impact Factor erschienen ist, sagt wenig über den tatsächlichen Citation Impact dieser Arbeit aus. Denn selbst in Zeitschriften mit einem hohen Journal Impact Factor hat die Masse der Arbeiten keinen ungewöhnlich hohen Citation Impact. Dem entsprechend wird eine typische (d. h. eine um den Median der Zitierungen liegende) Arbeit in einer Zeitschrift mit einem hohen Citation Impact nicht wesentlich häufiger zitiert als eine typische Arbeit in einer Zeitschrift mit einem niedrigen Citation Impact. Die nicht-symmetrische bzw. schiefe Verteilung der Zitierungen auf die Publikationen (Pareto-Verteilung) bedingt, dass nur ein kleiner Anteil vielzitatierter Arbeiten einen großen Anteil der Zitierungen erhält (*Seglen, 1992*). Die meistzitierte

Hälfte der Arbeiten wird um ein Vielfaches häufiger zitiert als der Rest. Der arithmetische Mittelwert (und damit der Journal Impact Factor einer Zeitschrift) wird nur von verhältnismäßig wenigen vielzitierten Arbeiten bestimmt, während die „typischeren“ Zitationszahlen erheblich niedriger liegen. Der Journal Impact Factor einer Zeitschrift erlaubt deshalb keinen Rückschluss auf den Citation Impact einer bestimmten in dieser Zeitschrift publizierten Arbeit. Ein extremes Beispiel ist die Zeitschrift *Acta Crystallographica A*, deren Journal Impact Factor innerhalb eines Jahres wegen einer einzelnen vielzitierten Arbeit von 2,051 (JCR-2008) auf 49,926 (JCR-2009) anstieg (Dimitrov, Kaveri, & Bayry, 2010).

Wir haben es hier also mit einer Korrelation in eine Richtung zu tun: Der Citation Impact von einzelnen Arbeiten verhilft einer Zeitschrift zu mehr oder weniger Ansehen; die Publikation einer Arbeit in einer Zeitschrift mit einem hohen Journal Impact Factor bewirkt jedoch nicht automatisch, dass diese auch hoch zitiert wird. Das Erscheinen einer Arbeit in einer Zeitschrift mit einem hohen Journal Impact Factor erhöht allenfalls die Sichtbarkeit der Arbeit und belegt zumeist, dass ein verhältnismäßig strenges Begutachtungsverfahren erfolgreich durchlaufen wurde. Das alles sind jedoch lediglich einige von vielen Voraussetzungen für die Entfaltung von Wirkung (Bornmann & Marx, *in press*). Deshalb ist es nicht verwunderlich, dass die Korrelation zwischen den Zitierungen der Arbeiten eines einzelnen Wissenschaftlers und den Journal Impact Factors der Zeitschriften dieser Arbeiten schwach ausfällt (Seglen, 1997). Journal Impact Factors sind zur Feststellung des Citation Impact der einzelnen Arbeiten eines Wissenschaftlers ungeeignet. Stattdessen sollte man die Zitierungen jeder einzelnen Publikation recherchieren. Der Umweg über die Journal Impact Factors macht keinen Sinn, auch wenn in einzelnen Disziplinen (wie z. B. der Medizin) oder in bestimmten Ländern (z. B. in skandinavischen Ländern, der Volksrepublik China und Indien) die Forschungsbewertung und Mittelvergabe gerne danach bemessen wird. Auch von den diversen in Gebrauch gekommenen Formeln zur Berechnung eines sogenannten „Personal Impact Factors“ möchten wir abraten.

7.2 Alternative Indikatoren

Grundlage der quantitativen (bibliometrischen) Bewertung von Wissenschaftlern und Forschungseinrichtungen sollten in jedem Fall die Zitationszahlen der einzelnen zu bewertenden Arbeiten sein. Für diesen Zweck wurde inzwischen eine Vielzahl von bibliometrischen Indikatoren entwickelt.² In den vergangenen Jahren entwickelte sich der h index neben dem Journal Impact Factor zum bekanntesten bibliometrischen Indikator (Bornmann & Marx, 2011). Der h index ist wie folgt definiert: „A scientist has

² Die Aussagekraft und die Grenzen quantitativer (bibliometrischer) Indikatoren zur Forschungsbewertung wurden wiederholt ausführlich diskutiert und werden deshalb hier nur grundsätzlich und kurz dargelegt (Anon, 2006, 2010; Bornmann & Daniel, 2008; Marx, 2009, 2011; Moed, 2005; van Noorden, 2010).

index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each" (Hirsch, 2005, S. 16569). Die Möglichkeit der einfachen und schnellen Berechnung sowie die Reduktion von Forschungsleistung auf eine einzige Zahl haben seinen Bekanntheitsgrad sehr begünstigt (Bornmann & Daniel, 2007). Der h index misst die Produktivität (Anzahl der Publikationen) und die Wirkung (Anzahl der Zitierungen) in einer einzigen Maßzahl. Allerdings ist der h index (wie viele andere Indikatoren auch) vom Schaffensalter und Fachgebiet abhängig, so dass nur Wissenschaftler mit vergleichbarem Alter und gleicher Fachdisziplin miteinander verglichen werden können.

Der h index, der ursprünglich für die Bewertung von Wissenschaftlern entwickelt wurde, kann grundsätzlich auf jedes Publikationsensemble angewendet werden, sofern die Zitationszahlen der einzelnen Publikationen im Set bekannt sind. Deshalb bietet sich damit die Möglichkeit, auch Zeitschriften zu bewerten. Braun, Glänzel und Schubert (2006) schlagen einen „ h -type index“ für Zeitschriften als Ergänzung des Journal Impact Factors vor. Einige Studien haben jedoch gezeigt, dass der h index von Zeitschriften hoch mit deren Journal Impact Factor korreliert und somit redundante Ergebnisse liefert (Bornmann, Marx, & Schier, 2009).

Aus der Notwendigkeit, Zitationszahlen fachlich und zeitlich zu normieren, ist die relative Zitationsrate als der zurzeit weit verbreitetste Indikator für den Vergleich einzelner Arbeiten (bzw. von Wissenschaftlern) hervorgegangen. Für die Berechnung der relativen Zitationsrate werden die tatsächlichen Zitierungen jeder einzelnen Arbeit (eines Wissenschaftlers, eines Forschungsinstituts usw.) den Erwartungswerten als Bezugsgröße gegenübergestellt, welche auf der Basis der durchschnittlichen Zitierungen aller im gleichen Jahr in der jeweiligen Zeitschrift erschienenen Arbeiten vom gleichen Dokumenttyp (d. h. Article oder Review) berechnet werden. Nimmt man eine einzelne Zeitschrift als Berechnungsgrundlage, erhält man zeitschriftenbezogene relative Zitationsraten. Alternativ kann man den Erwartungswert anstatt auf die einzelnen Zeitschriften auch auf Fachkategorien beziehen, die durch Zeitschriften-Sets definiert sind. Dadurch erhält man fachbezogene relative Zitationsraten. Ein relativer journal- oder fachbezogener Index von 1,2 bedeutet beispielsweise, dass die Wirkung der zu bewertenden Arbeiten um 20 Prozent über den Vergleichsarbeiten liegt. Inzwischen wurden auch Indikatoren entwickelt, welche die Zitate nicht gleichwertig zählen, sondern deren unterschiedliches Gewicht berücksichtigen: Je häufiger eine zitierende Arbeit ihrerseits zitiert wurde, desto stärker begünstigt sie die zitierte Arbeit.

8 Grundsätzliches zur Bibliometrie

8.1 Grenzen der Bibliometrie

Nur die Forscher selbst haben den für eine inhaltliche Bewertung von Forschungsarbeiten notwendigen Sachverstand. Forschungsleistung sollte deshalb niemals ausschließlich auf der Basis quantitativer Indikatoren, sondern immer nur mit Hilfe von Fachgutachtern bewertet werden (Informed Peer Review) (*Bornmann, 2011*). Die Bibliometrie ist also keinesfalls als Ersatz der Gutachter, sondern als zusätzliche quantitative Methode anzusehen. Allerdings bedarf die Erstellung und Interpretation bibliometrischer Daten einiger Erfahrung und Hintergrundinformationen, die nur Spezialisten in diesem Bereich haben. Der Wunsch nach verführerischen Rankings und Indikatoren wird durch einfach zu bedienende Werkzeuge erfüllt (siehe z. B. die Tools SciVal Spotlight von Elsevier und InCites von Thomson Reuters). Dieser wachsende Trend hin zu einer „Amateur-Bibliometrie“ stellt eine ernste Gefahr für die Forschungsevaluation dar.

Indem der zitierende Autor auf eine bestimmte Arbeit verweist, misst er dieser in der Regel eine gewisse Bedeutung bei. Die Bedeutung einer Arbeit stützt sich allerdings nicht nur auf einen, sondern auf mehrere Qualitätsaspekte. Qualität bezieht also ganz unterschiedliche Aspekte ein: Wirkung, Eleganz, Originalität, Wichtigkeit, Richtigkeit, aber auch Popularität oder gar Verwertbarkeit. Qualität ist also ein mehrdimensionales Phänomen, das dem entsprechend auch nur schwer quantitativ messbar ist. Zitierungen dokumentieren zunächst lediglich die Wirkung einer Arbeit in Form der Beachtung durch Fachkollegen. Da die anderen Aspekte durch Zitierungen kaum gemessen werden können, sind Zitierungen keine abschließenden Werturteile. Zahlreiche Untersuchungen deuten jedoch darauf hin, dass Wirkung (gemessen anhand von Zitierungen) und Wert bzw. Bedeutung (gemessen anhand der Einschätzung der Fachkollegen) korrelieren (*Bornmann, 2011*). So haben beispielsweise einige Studien zeigen können, dass Manuskripte, die von einer Zeitschrift zur Publikation angenommen wurden, im Durchschnitt häufiger zitiert werden als Manuskripte, die von der Zeitschrift abgelehnt und andernorts publiziert wurden. Indem Zitationszahlen die Wirkung als einen wichtigen Teilaspekt von Qualität quantifizieren, können sie durchaus als Stellvertreter (Proxy-Daten) für deren Wert aufgefasst und eingesetzt werden (*Martin & Irvine, 1983*).

8.2 Möglichkeiten der Verzerrung

Der sinnvolle Einsatz der Bibliometrie setzt hinreichend große Ensembles von zu untersuchenden Arbeiten voraus: Auf der Makroebene (ganze Länder) sind die Ergebnisse bibliometrischer Analysen statistisch am besten abgesichert; auch auf der Mesoebene (Universitäten, Forschungsinstitute, Forschungsgruppen) ist die Aussagekraft bibliometrischer Methoden hoch. Die Anwendung auf der Mikroebene (einzelne

Forscher, einzelne Arbeiten) erfordert dagegen eine vorsichtige Interpretation der Indikatoren. Auswertungen, die nicht von Fachleuten in der Bibliometrie vorgenommen werden, sind deshalb in der Regel problematisch bis irreführend. Die Aussagekraft von Peer Review verhält sich gegenüber der Bibliometrie eher umgekehrt: Die Aussagekraft von Gutachterurteilen ist auf der Mikroebene am größten (sofern die Gutachter einzelne Arbeiten wirklich lesen) und auf der Makroebene am geringsten.

Der Vorgang des Zitierens wird nicht ausschließlich von rein sachlichen Erwägungen bestimmt, sondern ist als sozialer Akt mit einem breiten Ermessensspielraum behaftet (*Bornmann & Daniel, 2008*). So zitieren beispielsweise Forscher gezielt Arbeiten von einflussreichen anderen Forschern in ihrem Fachgebiet, um die eigene Forschung zu untermauern und andere Forscher (die Leser) von der Richtigkeit und Wichtigkeit der eigenen Forschung zu überzeugen (*Gilbert, 1977*). Es ist unbestritten, dass dies zu Verzerrungen bei zitierungs-basierten Bewertungen führen kann. Die Verzerrungen wirken jedoch sowohl in Richtung einer Über- wie auch einer Unterbewertung. Inhaltlich überflüssige Zitate begünstigen z. B. den Citation Impact der zitierten Arbeiten, und absichtlich oder versehentlich unterlassene Zitierungen reduzieren ihn. Man kann deshalb davon ausgehen, dass sie sich bei der Untersuchung größerer Ensembles von Arbeiten herausmitteln und nur im Falle einzelner Arbeiten oder kleiner Ensembles eine Fehlerquelle darstellen.

In diesem Zusammenhang sollte jedoch nicht vergessen werden, dass auch die Forschungsbewertung durch Fachgutachter durchaus ihre Probleme mit sich bringt. So besteht z. B. die Gefahr, dass Gutachter die Bedeutung einer wissenschaftlichen Arbeit nicht erkennen oder sachfremde Einflüsse in die Beurteilung einfließen (*Bornmann, 2011*). Wenn es um die Einschätzung der Arbeit von Fachkollegen geht, sind auch Wissenschaftler nicht immer objektiv. Dies ist ein Grund, warum wiederholt objektive und nachvollziehbare Kriterien der Bewertung verlangt wurden, welche die Urteile der Fachgutachter ergänzen können. Eine auf Zitierungen basierende Methode bietet zwei entscheidende Vorteile, weshalb sie eine wichtige Unterstützung der Peers darstellt: Sie beruht einerseits auf einem größeren Blickwinkel, indem sie das Votum vieler Fachkollegen heranzieht, und ihre Ergebnisse sind andererseits überprüfbar. Abwägend und vernünftig eingesetzt, liefert die Bibliometrie durchaus Hinweise auf starke oder schwache Forschungsleistungen.

9 Diskussion

Journal Impact Factors wurden ursprünglich zur bibliometrischen Messung der Wirkung von Zeitschriften als Ganzes erfunden und eingesetzt. Unter Berücksichtigung der oben aufgeführten Einschränkungen sind Journal Impact Factors durchaus dafür geeignet, das Ansehen bzw. den Einfluss von Zeitschriften innerhalb ihrer jeweiligen

Fachdisziplin abzuschätzen. Doch selbst der Datenbankhersteller Thomson Reuters rät, den Journal Impact Factor vorsichtig und nur in Kombination mit anderen Kriterien und Indikatoren heranzuziehen. Die Verwendung von Journal Impact Factors zur Evaluierung von Wissenschaftlern stellt demgegenüber einen Missbrauch dieses Indikators dar. Der Einsatz der Journal Impact Factors in diesem Bereich wird durch die leichte Verfügbarkeit der Journal Impact Factors, einen vermeintlichen Mangel an Alternativen sowie ein fehlendes Verständnis bibliometrischer Grundzusammenhänge begünstigt.

Journal Impact Factors sind keine Maßzahl für die Qualität einzelner Arbeiten bzw. der Arbeiten eines Wissenschaftlers oder einer Forschungseinrichtung. Deshalb ist dieser Indikator für die Bewertung der Publikationen von Wissenschaftlern ungeeignet (und wegen erheblich besserer Alternativen inzwischen auch überflüssig). Grundlage der quantitativen (bibliometrischen) Bewertung von Wissenschaftlern und Forschungseinrichtungen sollten nicht die Journal Impact Factors der Zeitschriften sein, in denen die zu bewertenden Arbeiten erschienen sind, sondern die Zitationszahlen der einzelnen zu bewertenden Arbeiten. Auf dieser Basis wurde inzwischen eine Vielzahl von bibliometrischen Indikatoren entwickelt, von denen die fach- und zeitnormierten Indikatoren (relative Zitationsraten oder Perzentile) die problematischsten Verzerrungen am besten vermeiden (Bornmann, Mutz, Neuhaus, & Daniel, 2008; Leydesdorff, Bornmann, Mutz, & Opthof, 2011). Die Beschaffung von aussagekräftigen Zitationsdaten zu einzelnen Publikationen kann allerdings verhältnismäßig aufwendig und teuer sein.

Literatur

Adler, Robert; Ewing, John; Taylor, Peter (2009): Citation Statistics - A report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Statistical Science* 24(1), S. 1–14. doi: 10.1214/09-sts285.

Amin, Mayur; Mabe, Michael (2003): Impact factors: Use and abuse. *Medicina-Buenos Aires* 63(4), S. 347–354.

Editorial (2006): Cash-per-publication. *Nature* 441(7095), S. 785–786. doi: 10.1038/441786a.

Editorial (2010): How to improve the use of metrics. *Nature* 465(7300), S. 870–872. doi: 10.1038/465870a.

Archambault, Eric; Vignola-Gagne, Etienne (2004): Science-Metrix: The use of bibliometrics in the Social Sciences and Humanities. Montreal, Canada: Social Sciences and Humanities Research Council of Canada (SSHRCC).

Archambault, Eric; Vignola-Gagne, Etienne; Cote, Gregoire; Lariviere, Vincent; Gingras, Yves (2006): Benchmarking scientific output in the social sciences and humanities:

The limits of existing databases. *Scientometrics* 68(3), S. 329–342. doi: 10.1007/s11192-006-0115-z.

Bornmann, Lutz (2010): Mimicry in science? *Scientometrics* 86(1), S. 173–177. doi: 10.1007/s11192-010-0222-8.

Bornmann, Lutz (2011): Scientific peer review. *Annual Review of Information Science and Technology* 45, S. 199–245.

Bornmann, Lutz; Daniel, Hans-Dieter (2007): What do we know about the h index? *Journal of the American Society for Information Science and Technology* 58(9), S. 1381–1385. doi: 10.1002/asi.20609.

Bornmann, Lutz; Daniel, Hans-Dieter (2008): What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation* 64(1), S. 45–80. doi: 10.1108/00220410810844150.

Bornmann, Lutz; Marx, Werner (2011): The h index as a research performance indicator. *European Science Editing* 37(3), S. 77–80.

Bornmann, Lutz; Marx, Werner (in press): The Anna Karenina principle: A way of thinking about success in science. *Journal of the American Society of Information Science and Technology*.

Bornmann, Lutz; Marx, Werner; Gasparyan, Armen Yuri; Kitas, George D. (in press): Diversity, value and limitations of the Journal Impact Factor and alternative metrics. *Rheumatology International (Clinical and Experimental Investigations)*. doi: 10.1007/s00296-011-2276-1.

Bornmann, Lutz; Marx, Werner; Schier, Hermann (2009): Hirsch-type index values for organic chemistry journals: a comparison of new metrics with the Journal Impact Factor. *European Journal of Organic Chemistry* 2009(10), S. 1471–1476. doi: 10.1002/ejoc.200801243.

Bornmann, Lutz; Mutz, Rüdiger; Neuhaus, Christoph; Daniel, Hans-Dieter (2008): Use of citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics* 8, S. 93–102. doi: 10.3354/esep00084.

Bradford, Samuel C. (1934): Sources of information on specific subjects. *Engineering* 137, 85–86.

Braun, Tibor; Glänzel, Wolfgang; Schubert, Andras (2006): A Hirsch-type index for journals. *Scientometrics* 69(1), S. 169–173. doi: 10.1007/s11192-006-0147-4.

Dimitrov, Jordan D.; Kaveri, Srinivasa V.; Bayry, Jagadeesh (2010): Metrics: journal's impact factor skewed by a single paper. *Nature* 466(7303), S. 179. doi: 10.1038/466179b.

Garfield, Eugene (2006): The history and meaning of the Journal Impact Factor. *Journal of the American Medical Association* 295(1), S. 90–93. doi: 10.1001/jama.295.1.90.

Gilbert, G. Nigel (1977): Referencing as persuasion. *Social Studies of Science* 7(1), S. 113–122.

Glänzel, Wolfgang; Moed, Henk F. (2002): Journal impact measures in bibliometric research. *Scientometrics* 53(2), S. 171–193. doi: 10.1023/A:1014848323806.

Hirsch, Jorge E. (2005): An index to quantify an individual's scientific research output. Proceedings of the National Academy of Sciences of the United States of America 102(46), S. 16569–16572. doi: 10.1073/pnas.0507655102.

Kaltenborn, Karl-Franz; Kuhn, Klaus (2003): Der Impact-Faktor als Parameter zur Evaluation von Forscherinnen/Forschern und Forschung. Medizinische Klinik 98(3), S. 153–169. doi: 10.1007/s00063-003-1240-6.

Leydesdorff, Loet; Bornmann, Lutz; Mutz, Rüdiger; Opthof, Tobias (2011): Turning the tables in citation analysis one more time: principles for comparing sets of documents. Journal of the American Society for Information Science and Technology 62(7), S. 1370–1381. doi: 10.1002/asi.21534.

Martin, Ben R.; Irvine, John (1983): Assessing basic research - Some partial indicators of scientific progress in radio astronomy. Research Policy 12(2), S. 61–90.

Marx, Werner (2009): Forschungsbewertung auf der Basis von Zitierungen - Aussagekraft und Grenzen der Methode. In: Alexander von Humboldt-Stiftung (Hrsg.): Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen - Beiträge zur Beurteilung von Forschungsleistungen (S. 132–155). Alexander von Humboldt-Stiftung, Berlin, Germany.

Marx, Werner (2011): Bibliometrie in der Forschungsbewertung - Aussagekraft und Grenzen. Forschung & Lehre 18(11), S. 858–860.

Moed, Henk F. (2005): Citation analysis in research evaluation. Springer, Dordrecht, The Netherlands.

Reedijk, Jan (2011): Citations and ethics. Angewandte Chemie - International Edition 51(4), S. 828–830. doi: 10.1002/anie.201107554.

Seglen, Per O. (1992): The skewness of science. Journal of the American Society for Information Science 43(9), S. 628–638.

Seglen, Per O. (1997): Why the impact factor of journals should not be used for evaluating research British Medical Journal 314(7079), S. 498–502.

Thomson Reuters. (2012a): Essential Science Indicators. From http://thomsonreuters.com/products_services/science/science_products/a-z/essential_science_indicators/ (Zugriff: 17. März 2012)

Thomson Reuters. (2012b): Journal Citation Reports. From <http://scientific.thomson.com/products/jcr/> (Zugriff: 17. März 2012)

Van Noorden, Richard (2010): Metrics: a profusion of measures. Nature 465, S. 864–866. doi: 10.1038/465864a.

Vanclay, Jerome K. (in press): Impact Factor: Outdated artefact or stepping-stone to journal certification? Scientometrics.

Wilhite, Allen W.; Fong, Eric A. (2012): Coercive citation in academic publishing. Science 335(6068), S. 542–543. doi: 10.1126/science.1212540.

Anschrift der Verfasser:

Dr. Werner Marx
Max-Planck-Institut für Festkörperforschung
Informationsvermittlungsstelle (IVS-CPT)
Heisenbergstraße 1
70569 Stuttgart
E-Mail: w.marx@fkf.mpg.de

Dr. Dr. habil. Lutz Bornmann
Generalverwaltung der Max-Planck-Gesellschaft
Stabsreferat für Wissenschafts- und Innovationsforschung, Forschungsanalyse
Hofgartenstr. 8
80539 München
E-Mail: bornmann@gv.mpg.de
www.lutz-bornmann.de

Werner Marx arbeitet am Max-Planck-Institut für Festkörperforschung in Stuttgart. Er ist verantwortlich für die Servicegruppe Fachinformation, einer Einrichtung zur Unterstützung der Max-Planck-Institute der Chemisch-Physikalisch-Technischen Sektion der Max-Planck-Gesellschaft in Fragen der wissenschaftlichen Information und Forschungsbewertung.

Lutz Bornmann arbeitet als Wissenschaftssoziologe im Stabsreferat für Wissenschafts- und Innovationsforschung, Forschungsanalyse in der Generalverwaltung der Max-Planck-Gesellschaft. Er ist dort für Fragen der Forschungsevaluation und Bibliometrie zuständig.

Zitationen als Grundlage von Forschungsleistungsrankings – Konzeptionelle Überlegungen am Beispiel der Betriebswirtschaftslehre¹

Harry Müller

Dieser Beitrag widmet sich den Problemen und Anwendungsfeldern der Forschungsleistungsmessung mit Hilfe von Zitationen. Dazu ist zunächst grundlegend zu klären, weshalb Zitationen als Indikator für Forschungsleistung geeignet sind und wo die Grenzen dieses Konzepts liegen. Darauf aufbauend werden die gängigen Methoden zur Erstellung wissenschaftlicher Rankings vorgestellt und kritisch bewertet. Dabei stellt sich heraus, dass für eine deutschsprachige Sozialwissenschaft wie die Betriebswirtschaftslehre, an deren Beispiel in diesem Beitrag argumentiert wird, zahlreiche Besonderheiten zu berücksichtigen sind. Bei der Wahl der geeigneten Rankingmethode muss bedacht werden, dass die Publikationen größtenteils in deutscher Sprache erscheinen und neben wissenschaftlichen Zeitschriften auch Monographien und Sammelbände wichtige Publikationsmedien sind. Bei Berücksichtigung dessen lässt sich folgern, dass zitationsbasierte Verfahren trotz ihrer spezifischen Schwächen die für viele Zwecke am besten geeignete Methodik darstellen.

1 Einleitung

Die Definition und die Operationalisierung von Zielgrößen werden im Rahmen des Hochschulmanagements zunehmend wichtiger. Ein zweckmäßiges Management setzt voraus, dass Outputgrößen bekannt sind, die dann unter Berücksichtigung der zu ihrer Erstellung nötigen Ressourcen bewertet und verglichen werden können. Fehlt die Information über den Output, kann weder die Effektivität einer Aktivität bewertet, noch das Effizienzziel sinnvoll angestrebt werden. Es ist dabei eine Besonderheit des Hochschulwesens, dass sowohl Forschung als auch Lehre komplexe Dienstleistungen darstellen, deren Output wesentlich schwieriger zu erfassen ist als die bei ihrer Erbringung verwendeten Ressourcen. Seit etwa zwei Jahrzehnten besteht seitens der Politik als Träger der (öffentlichen) Hochschulen, aber zunehmend auch seitens der Studierenden bzw. der Gesellschaft generell ein gesteigertes Interesse an einer Quantifizierung. Die Exzellenzinitiative des Bundes, Zielvereinbarungen bei der Budgetierung, leistungsbezogene Arbeitsverträge für Hochschullehrer und nicht zuletzt der verstärkte Einsatz von Lehrevaluationen zeugen von dieser Entwicklung (vgl. Hennig-Thurau/Walsh/Schrader 2004).

¹ Ich danke Alexander Dilger und einem anonymen Gutachter für wertvolle Anmerkungen und Kommentare.

Die Quantifizierung von Forschungsleistungen wird sowohl für die Hochschulen als auch für das an ihnen beschäftigte wissenschaftliche Personal zunehmend wichtiger. Die Forschungsleistungen der Beschäftigten sind viel stärker als die Lehrleistungen bestimmend für die Karriere an den Universitäten und das Ansehen in der *Scientific Community*. Angesichts der großen fachlichen Spezialisierung jeder einzelnen Forscherin und jedes einzelnen Forschers² ist das Aufstellen von adäquaten und intersubjektiv nachvollziehbaren Kriterien zur Beurteilung von Forschungsleistungen naturgemäß komplex. Vor diesem Hintergrund ist es nicht verwunderlich, dass die Eignung von bibliometrischen Verfahren als quantitatives Maß für die Bewertung einer wissenschaftlichen Arbeit umstritten ist und z. T. ganz grundsätzlich kritisiert wird (vgl. Kieser 2010). Einerseits können bibliometrische Methoden als ein transparentes und intersubjektiv nachvollziehbares Maß qualitative Verfahren wie z. B. *Peer Review* ergänzen. Kritiker wie Lawrence (2003) sehen jedoch im Bestreben, wissenschaftliche Leistungen messbar und damit auch für Außenstehende transparenter zu machen, eine Entwicklung zur „Audit Society“ (zum Begriff siehe Power 1997).

Dieser Einwand macht deutlich, dass bibliometrischen Verfahren als Managementinstrument an Hochschulen oder in der Wissenschaftspolitik eine große Bedeutung zukommt. Ihr Einsatz birgt Chancen und Risiken: Eine gesteigerte Transparenz hinsichtlich der Leistungserbringung innerhalb der Hochschulen wie auch gegenüber ihren Stakeholdern ist geeignet, die Qualität der Managemententscheidungen zu verbessern. Für ein Berufungsverfahren ergeben sich durch den Einsatz bibliometrischer Verfahren zusätzliche Informationen, anhand derer sich die Kandidaten differenzieren lassen. Zudem entsteht ein weiterer Anreiz zu qualitativ hochwertigen Veröffentlichungen, der z. B. durch die Verknüpfung mit Leistungszulagen im Rahmen der W-Besoldung weiter gestärkt werden kann (vgl. Blomeyer 2007; Süß 2007). In dieser Wirkung sehen Kritiker aber auch die Risiken im Einsatz der Bibliometrie zu Managementzwecken. Wissenschaftler seien oftmals stark intrinsisch motiviert, und bei der Implementierung von starken extrinsischen Anreizen, beispielsweise in Form von Leistungszulagen, bestehe die Gefahr der Verdrängung der intrinsischen Motivation (vgl. Frey 2003). Wenn deshalb Wissenschaftler ein stärkeres Interesse an Rangplätzen in Rankings als an der wissenschaftlichen Erkenntnis selbst entwickelten, hätten Rankings der akademischen Kultur schweren Schaden zugefügt (vgl. Weingart 2005).

² Im Interesse der Lesbarkeit wird bei Personenbezeichnungen teilweise nur die männliche Form verwendet. Hier sind immer auch weibliche Personen gemeint, sofern kein entgegenstehender Hinweis erfolgt.

2 Forschungsleistung und Zitationen

2.1 Forschungsleistung an Hochschulen

An den tertiären Bildungseinrichtungen im deutschsprachigen Raum nimmt die Forschung eine zentrale Rolle ein. Das Hochschulrahmengesetz (HRG) definiert sie als zentrale Aufgabe von Hochschulen: „Die Hochschulen dienen [...] der Pflege und Entwicklung der Wissenschaften und der Künste durch Forschung, Lehre, Studium und Weiterbildung“ (§ 2 HRG). Daneben wird Forschung nicht nur von Hochschulen, sondern auch von spezialisierten Forschungsinstituten (z. B. der Max-Planck-Gesellschaft) betrieben. Auch außerhalb von spezifischen Institutionen können Forschungsleistungen entstehen, wenn z. B. ein Privatgelehrter oder ein Unternehmen wissenschaftliche Erkenntnisse gewinnt.

Forschung wird im HRG als „Gewinnung wissenschaftlicher Erkenntnisse“ (§ 22 Abs. 1 HRG) definiert. Aus der Perspektive der Betriebswirtschaftslehre (BWL) lässt sich Forschung an Universitäten mit Dyckhoff et al. (2005) als „Produktion neuen öffentlichen Wissens über die Welt“ bezeichnen. Diese Definition weist auf zwei wesentliche Aspekte hin, die für den hier diskutierten Kontext relevant sind: Erstens kann die Forschungstätigkeit als eine Art Produktionsprozess und damit als eine Beziehung von Input und Output betrachtet werden. Bei der Forschung werden Ressourcen (Arbeitszeit und Kapital) verbraucht, und am Ende steht ein bestimmter Output, nämlich neues öffentliches Wissen. Wie dieser Output exakt aussieht und mit welchen Indikatoren er quantitativ und qualitativ gemessen werden kann, ist eine komplexe Fragestellung und Gegenstand dieses Beitrags. Zweitens sei auf den öffentlichen Charakter des Wissens hingewiesen, d. h. Wissen muss im Sinne dieser Definition nicht nur gewonnen, sondern auch publiziert werden. Aus dieser Perspektive betrachtet ist Forschung immer auch eine Form von Kommunikation und eine Idee, die niemanden mitgeteilt wird, jedoch noch kein vollwertiger Beitrag dazu.

Nicht öffentliche Erkenntnisse, wie sie beispielsweise in der Auftragsforschung für Industrie und Militär gewonnen werden, fallen somit nicht unter die Forschungsdefinition von Dyckhoff et al., obwohl sie mit wissenschaftlichen Methoden gewonnen sein und somit Forschungscharakter haben können. Ihnen fehlt jedoch die Öffentlichkeit, d. h. die Erkenntnisse können weder in der *Scientific Community* diskutiert werden, noch den anderen Forschern Möglichkeiten zur Erkenntnis bieten.³ Nicht veröffent-

³ Auch industrielle Auftragsforschung kann durch eine Patentierung öffentlich gemacht werden. Sie wird dadurch jedoch nicht zu einem öffentlichen Gut im Sinne der Ökonomie, weil ihre Verwendung nicht frei ist, sondern bezahlt werden muss. Dadurch wird die Messung und Bewertung privater Forschungsleistungen allerdings einfach, weil dieses Wissen monetär bewertet werden kann. Dies geschieht bei privatwirtschaftlichen Unternehmen gemäß den Rechnungslegungsvorschriften des IAS 38 (siehe Pellens u. a. 2008). Auch aus diesem Grund ist eine Abgrenzung von privatem und öffentlichem Wissen bei der Messung und Bewertung von Forschungsleistungen sinnvoll. Die im Rahmen dieses Beitrags diskutierten Probleme betreffen im Wesentlichen nämlich nur letzteres.

lichte Forschungsergebnisse der Auftragsforschung stellen daher keine positive Externalität im Sinne der mikroökonomischen Theorie dar. Sie sind somit kein öffentliches, sondern ein privates Gut (vgl. *grundlegend Musgrave/Musgrave 1976*). Während sich der Wert der privaten Auftragsforschung unmittelbar aus der Zahlungsbereitschaft des Auftraggebers ergibt, wird die öffentliche Forschung typischerweise staatlich finanziert und die Forschenden streben nicht nur nach einer materiellen, sondern in der Regel auch nach einer ideellen Entlohnung, die sich aus der Reputation in- und außerhalb des Fachs ergibt. Auch aus der mikroökonomischen Theorie der öffentlichen Güter lässt sich also eine Differenzierung zwischen veröffentlichten und privaten Forschungsergebnissen ableiten.

2.2 Grundfragen der Messbarkeit

Alle Ansätze, Forschungsleistungen verschiedener Wissenschaftler miteinander zu vergleichen, sind mit spezifischen Problemen und Schwierigkeiten verbunden. Das Quantifizieren wissenschaftlicher Exzellenz mit Hilfe einer Kennzahl oder eines Rangplatzes ist in der Regel nicht unkritisch (vgl. *Adler/Harzing 2009*). Die Ergebnisse werden dabei oft maßgeblich von der verwendeten Methodik beeinflusst. Rein qualitative Begutachtungsansätze wie z. B. die *Peer Review* können diese Problematik aber ebenfalls nicht zufriedenstellend lösen, da sie von der Position des Gutachters abhängen und andere Gutachter typischerweise zu anderen Ergebnissen kommen. Die intersubjektive Akzeptanz solcher Ergebnisse dürfte aufgrund der im Bewertungsprozess zwangsläufig inhärenten Intransparenz tendenziell eher niedrig sein.⁴ Nicht zuletzt sind Vergleiche auf der Grundlage einer *Peer Review* lediglich für zahlenmäßig kleine Beobachtungseinheiten praktikabel und für die Betrachtung einer größeren Zahl von Personen und Veröffentlichungen nur unter Inkaufnahme eines (ggf. prohibitiv) hohen Aufwands durchführbar. Das gilt vor allem dann, wenn man sich nicht nur auf Zeitschriftenartikel beschränkt, deren Begutachtung bereits im Einreichungsverfahren stattgefunden hat, sondern auch Monographien mit einschließen möchte.

Die quantitativen Ansätze zur Messung von Forschungsleistungen gehen davon aus, dass wissenschaftlicher Wettbewerb im Kern ein Wettbewerb um Reputation ist, in dem die Forscher danach streben, neue Erkenntnisse zu gewinnen und diese in der *Scientific Community* zu verbreiten (vgl. *Klingemann 1988; sowie Dyckhoff/Schmitz 2007*). Angeknüpft wird dabei ausdrücklich an der Veröffentlichung einer Erkenntnis und nicht bereits an der Erkenntnis selbst. Eine bedeutsame Entdeckung, die nicht an das (Fach-)Publikum kommuniziert wird, dürfte kaum zum wissenschaftlichen Fortschritt beitragen. Wird die Erkenntnis auf eine geeignete Art publiziert, so trägt sie in

⁴ Transparent wäre das Verfahren allenfalls bei einer Veröffentlichung der Gutachten, was bei wissenschaftlichen Zeitschriften allerdings nicht üblich ist und generell im Bezug auf die Autonomie der Gutachter problematisch sein dürfte.

der *Scientific Community* zu weiteren neuen Entdeckungen bei, da die Forscherkollegen an die Erkenntnisse anknüpfen können.

Die Nutzung bibliometrischer Methoden zur Messung von Forschungsleistungen hat in den deutschsprachigen Sozialwissenschaften noch keine lange Tradition, sondern ist ein vergleichsweise neues Konzept. Dabei konnte man sich auf Pionierarbeiten aus dem englischsprachigen Ausland und aus anderen Wissenschaften beziehen, wobei die Entwicklung von Zitationsdatenbanken und Methoden zu ihrer Auswertung in den 1960er und 1970er Jahren vornehmlich von den Naturwissenschaften ausging (vgl. Osareh 1996). In diesen Disziplinen hat die Bewertung von Fachzeitschriften, Forschungsinstituten oder individuellen wissenschaftlichen Leistungen anhand bibliometrischer Indikatoren bereits eine längere Tradition, was zum Teil darauf zurückzuführen ist, dass dort nahezu ausnahmslos in englischer Sprache und in internationalen Fachzeitschriften publiziert wird (vgl. Henrekson/Waldenström 2008). Auch wenn erst seit der Entwicklung von Computersystemen bibliometrische Studien in größerem Umfang möglich sind, so ist die Idee, wissenschaftliche Literatur mit statistischen Methoden zu beschreiben, deutlich älter, wie frühe Studien von Campbell (1896) und Cole/Eales (1917) belegen.

2.3 Zitationen als Indikator für Forschungsleistung

Eine Grundregel des wissenschaftlichen Arbeitens besteht darin, die Nutzung fremden Gedankenguts durch einen entsprechenden Verweis darauf kenntlich zu machen. Die Autoren eines wissenschaftlichen Beitrags zitieren vornehmlich diejenigen Quellen, die ihnen einen Erkenntnisgewinn gebracht und ihre eigenen Ideen beeinflusst haben. Die Zitation spiegelt auf diese Weise die Anerkennung einer wissenschaftlichen (Vor-)Arbeit wider (vgl. Klingemann 1988; Schlinghoff/Backes-Gellner 2002). Umgekehrt deuten wenige oder völlig ausbleibende Zitationen darauf hin, dass eine Veröffentlichung aus der Sicht der *Scientific Community* keinen wesentlichen Beitrag zum Erkenntnisfortschritt geleistet hat und als eher unbedeutend angesehen werden kann (vgl. Schmitz 2008). Auf diese Weise kann die Anzahl der Zitationen als Indikator für die Qualität einer wissenschaftlichen Veröffentlichung angesehen werden (vgl. Weingart/Winterhager 1984), zumindest jedoch als Indikator für die Rezeption einer Veröffentlichung (vgl. Schmitz 2008).

Diese Sichtweise ist allerdings nicht unumstritten. MacRoberts/MacRoberts (1996, S. 442) kommen zu folgendem Urteil:

„The basic assumptions of citation analysis are clearly false. The assumption ‘that the research cited by scientists in their own papers represents a roughly valid indicator of influence on their work’ has been falsified repeatedly, as has the assumption on which

it is based: that scientists are motivated to cite their influence or to give credit where credit is due.“

Dieser vernichtenden Kritik muss allerdings entgegengehalten werden, dass ein bloßes Falsifizieren der Annahmen durch die Präsentation eines geeigneten Gegenbeispiels (*siehe hierzu Popper 2005*) zur Klärung der Frage, ob Zitationen ein geeigneter Indikator für Forschungsleistung sind, wenig beiträgt. Entscheidend ist nicht, ob es einzelne Fälle gibt, die dem oben Erläuterten widersprechen, sondern ob die getroffenen Annahmen die wissenschaftliche Praxis adäquat beschreiben.

Ähnliches gilt für das in der Literatur vielfach problematisierte Phänomen der negativen Zitationen. So können wissenschaftliche Arbeiten auch deswegen zitiert werden, um auf Fehler in ihnen hinzuweisen. Fraglich ist allerdings, ob dieser Umstand die Validität von Zitationsindikatoren zur Forschungsleistungsmessung tatsächlich einschränkt. Thesen, die sich als falsch erwiesen haben, werden schließlich nur dann zitiert, wenn sie trotzdem als hinreichend bedeutsam anerkannt wurden und den neuen Erkenntnissen im Weg standen. Unbedeutendes dagegen wird ignoriert (*vgl. Meadows 1974*). Davon abgesehen werden revolutionär neue Ansätze in der Wissenschaft typischerweise anfangs sehr scharf kritisiert, bevor sie sich durchsetzen, und ehemals wirkmächtige, aber später widerlegte Theorien haben ebenso zur Fortentwicklung des Fachs beigetragen (*vgl. Garfield 1979*). Letztlich ist es auch nicht Aufgabe bibliometrischer Indikatoren, den Wahrheitsgehalt einer Publikation abzubilden, sondern den Einfluss einer Veröffentlichung auf die Kommunikation des Fachs zu messen. Das gilt insbesondere für Sozialwissenschaften wie die BWL, in denen ein analytischer Wahrheitsbegriff ohnehin kaum anzuwenden sein dürfte.

Empirische Studien bestätigen die hier angestellten Überlegungen zur Funktion einer wissenschaftlichen Zitation recht deutlich. So kommt Cronin (*1982*) bei einer Befragung im Fach Psychologie zu dem Ergebnis, dass 95 Prozent der befragten Forscher der Aussage zustimmen, dass Anerkennung in der *Scientific Community* über das Zitieren verteilt wird. Auch andere empirische Untersuchungen weisen in dieselbe Richtung (für einen Überblick siehe *Schmitz 2008*). Die überwiegende Mehrheit der Wissenschaftler versteht ihre Arbeitspraktiken so, dass ein Zusammenhang von Zitation und Anerkennung einer Leistung bzw. der Qualität einer Veröffentlichung gegeben ist (*vgl. Le Pair 1988*). Die Tatsache, dass es in einzelnen Fällen anders sein mag, dürfte die Validität zitationsbasierter Verfahren zur Forschungsleistungsmessung nur unwesentlich relativieren.

3 Rankings im Hochschulwesen

3.1 Rankings und die Kritik daran

Dilger (2010, S. 92) definiert Rankings als eine Methodik, bei der „Beobachtungseinheiten nach einem Kriterium in eine Reihenfolge gebracht“ werden. Diese Definition ist breit gefasst, denn es lassen sich vielerlei Dinge gemäß verschiedenen Kriterien sortieren. Beobachtungseinheit und Rangfolgenkriterium müssen also im Hinblick auf die Ziele des Rankings definiert werden. Im hier betrachteten Kontext soll die Messung von Forschungsleistungen im Mittelpunkt stehen. Beobachtungseinheiten sind daher diejenigen, die ebensolche Leistungen erbringen, namentlich die Forscher sowie im Aggregat die wirtschaftswissenschaftlichen Fachbereiche der Universitäten. Welches das zu diesem Zweck geeignete Rankingkriterium ist, soll im Folgenden diskutiert werden. Gebräuchlich sind vor allem gewichtete Zeitschriftenveröffentlichungen und Zitationskennzahlen.

Auch die Rankingziele sind umstritten, wie sich besonders in der Diskussion um die Handelsblatt-Rankings zeigte, in der neben methodologischen Anmerkungen (vgl. Hofmeister/Ursprung 2008 für das VWL- sowie Müller 2010 für das BWL-Ranking) auch ganz grundsätzliche Kritik laut wurde (vgl. Kieser 2010). Aus der letztgenannten Position folgt ein bewusster Verzicht auf die Informationen aus quantitativen Rankingmethoden. In bestimmten Situationen müssen Rangfolgeentscheidungen jedoch zwangsläufig getroffen werden, etwa in einem Berufungsverfahren. Lässt man hier quantitative Informationen bewusst außen vor, führt dies allenfalls zu einem intransparenten und subjektiven Ranking, das methodologisch nicht mehr kritisiert werden kann, weil das Verfahren der Rangfolgenbildung nicht expliziert wird. Sachgerechter erscheint es hier, Rankings „relativ zueinander und relativ zu anderen realistischen Alternativen“ (Dilger 2010, S. 92) zu diskutieren.

Gemäß dieser Argumentation erscheint eine differenzierte Anwendung von Rankingverfahren in einer konkreten Entscheidungssituation sinnvoll. Dabei können die Ergebnisse immer auch vor dem Hintergrund der methodologischen Schwächen der jeweils verwendeten Rankingmethode diskutiert werden. Personenrankings mit vielen Beobachtungseinheiten, wie z. B. das Handelsblatt-BWL-Ranking (Müller/Storbeck 2009) oder auch die Rankings von Dilger (2010) und Dilger/Müller (2012), sollten nicht unreflektiert im Sinne einer „akademischen Hitparade“ gelesen und interpretiert werden. Abgesehen von den Informationen über die namentlich genannten Bestplatzierten können nämlich auf diese Weise verschiedene Rankingmethoden exemplifiziert und kritisch diskutiert werden.

3.2 Methoden für Personenrankings

3.2.1 Zeitschriftenbasierte Personenrankings

Die zeitschriftenbasierte Methode zur Erstellung von Personenrankings ist in der BWL am meisten verbreitet und beruht auf einer Qualitätsgewichtung von Zeitschriften. Die jeweiligen Publikationen werden danach beurteilt, in welchen wissenschaftlichen Zeitschriften sie erschienen sind. Das impliziert, dass jede Zeitschrift einen gleichbleibenden Qualitätsstandard hat, der sich u. a. in der Annahme bzw. der Ablehnung von Einreichungen ausdrückt (vgl. *Hennig-Thurau/Walsh/Schrader 2004*). Das Zeitschriftenranking, das eine möglichst breite Abdeckung aufweisen sollte, wird auf diese Weise zur notwendigen Grundlage für ein darauf aufbauendes Personenranking.

Die zugrundeliegende Zeitschriftenliste kann auf verschiedene Bewertungsmethoden zur Ermittlung des Qualitätskriteriums zurückgreifen. Es kann z. B. auf bibliometrischen Daten oder auch auf Befragungsergebnissen basieren. Bibliometrische Zeitschriftenrankings beruhen in der Regel auf *Impact*-Faktoren, mit denen die durchschnittliche Zitationshäufigkeit der in den jeweiligen Zeitschriften erschienenen Beiträge abgebildet wird. Ein Beispiel hierfür ist die Zeitschriftenliste des *Social Sciences Citation Index* (SSCI). Das VHB-Jourqual ist in Deutschland das bekannteste befragungsbasierte Zeitschriftenranking, für das die Mitglieder des Verbands der Hochschullehrer für Betriebswirtschaft jeweils die Zeitschriften ihrer Teildisziplin einschätzen sollten (vgl. *Hennig-Thurau/Walsh/Schrader 2004*).

Ein entscheidender Vorteil zeitschriftenbasierter Personenrankings liegt in der vergleichsweise unproblematischen Erstellung, wenn bereits ein Zeitschriftenranking vorliegt. Ausgehend von den Qualitätskennzahlen der jeweiligen Zeitschriften können die Publikationen der Wissenschaftler mit Gewichten versehen und aufaddiert werden. Dieses Verfahren ist transparent und für den einzelnen Wissenschaftler ex ante gut abschätzbar. Aus der Qualitätsgewichtung der Zeitschriften, die als Qualitätsindikator auf die einzelnen Veröffentlichungen darin übertragen wird, ergibt sich ein unmittelbarer Anreiz, in hochrangigen Zeitschriften zu publizieren.

Ein diskretionärer Spielraum besteht dann, wenn mehrere Zeitschriftenrankings miteinander kombiniert werden, die auf jeweils unterschiedlichen Grundlagen basieren und eine unterschiedlich breite Abdeckung aufweisen. Auf diese Problematik ist bereits an anderer Stelle in Bezug auf das Handelsblatt-BWL-Ranking hingewiesen worden (vgl. *Müller 2010*).

Zeitschriftenbasierte Personenrankings sind allerdings konstruktionsbedingt mit einer Reihe von Problemen behaftet. Erstens können nur in Zeitschriften publizierte Forschungsbeiträge für das Ranking verwendet werden. Monographien, Beiträge in

Sammelbänden oder Online-Publikationen bleiben methodenbedingt unberücksichtigt. Weil auch in diesen Medien für die Fortentwicklung der BWL bedeutende Forschungsergebnisse veröffentlicht werden (vgl. *Griffith/Tamer Cavusgil/Xu 2008; Neuhaus/Daniel 2008*), ist die exklusive Betrachtung von Zeitschriftenbeiträgen inhaltlich nicht zu rechtfertigen. Sie wird in der Regel damit verteidigt, dass für andere Publikationstypen vergleichbare Qualitätsindikatoren fehlen (vgl. *Hennig-Thurau/Walsh/Schrader 2004, S. 521*). Diese Argumentation verliert allerdings dann an Gewicht, wenn die Messung von Qualitätsindikatoren (z. B. durch Zitationen) auf alle Publikationsarten ausgedehnt werden kann. Dies ist heute mit umfassenden Datenbanken wie z. B. *Google Scholar* durchaus möglich.

Zweitens bedeutet die Bewertung eines Forschungsbeitrags danach, in welchem Medium er erschienen ist, immer die Zuweisung eines Mittelwerts. Der Wert der Zeitschrift wird durch die Zitationen bestimmt, die auf die in ihr publizierten Beiträge entfallen. Der so im Zeitschriftenranking gemittelte *Impact*-Faktor wird nun wieder den einzelnen Beiträgen zugeordnet. Auf diese Weise werden die verfügbaren Informationen nicht bestmöglich ausgenutzt. Zeitschriftenbasierte Personenrankings enthalten eine unnötige Unschärfe an dieser Stelle. Wirklich herausragende Publikationen werden auf diese Weise tendenziell unterschätzt. Seglen (1994) hat nachgewiesen, dass der *Impact*-Faktor einer Zeitschrift nur sehr schwach mit dem *Impact* der darin enthaltenen Beiträge korreliert ist. Die Zitationen der Artikel bilden das Renommee der Zeitschrift, aber nicht umgekehrt (siehe auch Seglen 1997). So sind selbst in erst-rangigen Zeitschriften viele Beiträge zu finden, die nur sehr wenig oder überhaupt nicht rezipiert wurden, während für die weitere Forschung bedeutende Erkenntnisse mitunter auch in zweitrangigen Journals publiziert wurden (vgl. *Frey/Rost 2010*).

Im Fall von Personenrankings, die auf bibliometrisch erzeugten Zeitschriftenrankings basieren, stellen sich darüber hinaus weitere Probleme. Weisen die Zitationsdatenbanken größere systematische Lücken auf, können daraus Verzerrungen auch auf der Ebene des Personenrankings resultieren. Deutlich wird das, wenn man sich z. B. die *Impact*-Faktoren der Zeitschriften im SSCI ansieht. Der SSCI enthält kaum deutschsprachige Zeitschriften der BWL, die *Betriebswirtschaftliche Forschung und Praxis* (BFuP) ist eine der wenigen Ausnahmen. Clermont/Schmitz (2008, S. 23) folgern daraus, dass es mit dem SSCI „nicht möglich [sei], deutschsprachig publizierte Forschungsbeiträge oder ihre Zitationen [...] sinnvoll zu erheben“. Weil wissenschaftliche Diskurse darüber hinaus oft regionale Schwerpunkte haben, werden auch die *Impact*-Faktoren der wenigen im SSCI enthaltenen deutschsprachigen Zeitschriften niedriger ausgewiesen, als sie in einer vollständigen Datenbank wären. Da deutschsprachige Beiträge vornehmlich von deutschsprachigen Beiträgen zitiert werden, entsteht eine systematische Verzerrung zulasten der deutschsprachigen Literatur. So fehlen den Beiträgen in der BFuP beispielsweise die Zitationen aus der nicht im SSCI enthaltenen

Zeitschrift für Betriebswirtschaft (ZfB). Auf diese Weise erscheinen auf Deutsch publizierte Beiträge weniger wirkmächtig als englischsprachige.⁵

Die Kombination mehrerer unvollständiger Zeitschriftenrankings zu einer gemeinsamen Liste (für einen *methodologisch anspruchsvollen Ansatz* vgl. *Schulze/Warming/Wiermann 2008*) zwecks Erstellung eines Personenrankings löst das Problem nur unvollständig (vgl. *Albers 2009 zur Kritik*). Zwar entsteht auf diese Weise die gewünschte Abdeckungsbreite. Die systematische Verzerrung zu Ungunsten der nicht-englischsprachigen Literatur verschwindet jedoch nicht. Werden die deutschen Titel in jedem Ranking systematisch unterhalb ihres eigentlichen (d. h. bei einer vollständigen Datenbank sich ergebenden) Werts aufgeführt, so werden sie auch in der aggregierten Zeitschriftenliste nicht auf adäquaten Plätzen liegen.

Darüber hinaus können sich die dem Personenranking zugrunde liegenden Zeitschriftenrankings im Laufe der Zeit verändern. *Impact*-Faktoren sind typischerweise nicht konstant, sondern zeichnen die Entwicklung einer Zeitschrift nach. Da auch vor vielen Jahren erschienene Beiträge jeweils mit den aktuellen *Impact*-Faktoren (oder Befragungsdaten) der Zeitschriften bewertet werden, entstehen im Laufe der Zeit zufällige Effekte zu Gunsten oder zu Ungunsten der damaligen Autoren. Diese Auf- oder Abwertungen haben in Bezug auf die zu bewertenden alten Publikationen keine reale Substanz: Wird beispielsweise eine Zeitschrift von VHB-Jourqual 1 zu Jourqual 2 heruntergestuft, so sagt dies wenig über einen vor zehn Jahren in ihr erschienenen Beitrag aus.⁶

Schließlich ist auch beim zeitschriftenbasierten Personenranking im schlimmsten Fall mit strategischem Zitieren zu rechnen (vgl. *Dilger 2000*). Die Autoren einer Zeitschrift haben ein Interesse daran, ebendiese in der Folgezeit häufig zu zitieren, um über den *Impact*-Faktor der Zeitschrift die Bewertung ihrer eigenen Beiträge zu verbessern. Aufgrund von Kollektivgutproblemen ist der Anreiz zur Selbstzitation bei zeitschriften-

⁵ Das schließt selbstverständlich nicht aus, dass deutschsprachige Beiträge tatsächlich weniger wirkmächtig als englischsprachige Publikationen sind. Ob dies jedoch so ist, ließe sich gerade durch eine Zitationsanalyse mit einer beide Sprachen abdeckenden Datenbank herausfinden. Im Rahmen eines umfassenden Rankings, in das deutsche und englischsprachige Veröffentlichungen einbezogen sind, wird sich eine größere Bedeutung von internationalen Publikationen in den Zitationen niederschlagen. Sich stattdessen bewusst nur auf englische Publikationen zu beschränken, weil diesen per se ein vermeintlich höherer Wert zugemessen wird, erscheint dagegen nicht sinnvoll. Bei der Verwendung von Zitationen als Maßstab ist die Gewichtung von deutschen gegenüber internationalen Publikationen nicht länger ein der Rankingmethode inhärentes Werturteil, sondern sie wird zu einer empirischen Frage.

⁶ Dieser Einwand könnte z. B. dadurch adressiert werden, dass jedem Beitrag der jeweils für das Jahr seiner Publikation gültige *Impact*-Faktor zugewiesen wird. Auf diese Weise erhöht sich jedoch der Erstellungsaufwand, weil zusätzlich zu den Zeitschriften auch noch zwischen ihren jeweiligen Jahrgängen unterschieden werden muss. Oftmals sind die Beiträge zudem älter als die frühesten verfügbaren Zeitschriftenrankings. In der Praxis ist daher eine Verwendung zeitpunktspezifischer Zeitschriftengewichte zur Erstellung von Personenrankings zumindest in der BWL bislang unüblich (*Müller/Storbeck 2009*; vgl. *Wolf/Rohn/Macharzina 2006*).

basierten Personenrankings allerdings geringer als bei zitationsbasierten. Zeitschriftenrankings auf der Basis von Expertenurteilen reizen ein ähnliches Verhalten an: Wissenschaftler sind möglicherweise dazu verleitet, die Zeitschriften mit ihren eigenen Beiträgen besonders gut zu bewerten (vgl. Breuer 2009). Aus diesem Grund kommt der Anzahl der für jede Zeitschrift abgegebenen Einschätzungen z. B. beim Jourqual 2 eine wichtige Bedeutung zu.

3.2.2 Zitationsbasierte Personenrankings

Wird die Zitation als Indikation für die Rezeption und Anerkennung einer Forschungsleistung verstanden, so kann anhand dieser Kennzahl ein Ranking konstruiert werden. Nicht der Umstand der Veröffentlichung bzw. die Platzierung eines Beitrags in einem bestimmten wissenschaftlichen Journal wird als Qualitätsgewicht herangezogen, sondern es wird direkt auf den *Impact* des einzelnen Beitrags abgestellt. Die Anzahl der Zitationen als Qualitätsindikator zu nutzen ist kongruent mit dem oben erläuterten Konzept des wissenschaftlichen Wettbewerbs. Qualität wird nicht durch das Urteil des Herausgebers einer Zeitschrift (und der ggf. von ihm bestellten Gutachter), sondern durch die Reaktion der gesamten *Scientific Community* abgebildet.

Weil die Zitationen direkt für jede einzelne Veröffentlichung jedes Forschers erhoben werden, entfällt die im Falle der zeitschriftenbasierten Rankings kritisierte Mittelwertbildung. Es wird nur der individuelle *Impact* jedes einzelnen Beitrags gemessen. Externe Effekte von anderen im gleichen Medium veröffentlichten Beiträgen existieren höchstens indirekt, wenn eine hochrangige Zeitschrift generell mehr Leser hat und ein Beitrag darin ein größeres Publikum und mehr potenzielle Rezipienten erreicht.

Ein zweiter wichtiger Vorteil, den zitationsbasierte Rankings gegenüber zeitschriftenbasierten haben, liegt im grundsätzlich breiteren Fokus der Betrachtung. Die Rankingmethode ist nicht auf Zeitschriftenbeiträge begrenzt, sondern kann auch Monographien und Beiträge in Sammelbänden umfassen, sofern die Zitationen dieser Veröffentlichungen erhoben werden können. Die Qualität des zitationsbasierten Rankings ist daher eng mit der Qualität und der Abdeckungsbreite der zugrundeliegenden Datenbank verknüpft. Der Marktführer im Bereich der wissenschaftlichen Datenbanken, das ISI *Web of Science* (SSCI, SCI-X und A&HCI), erfüllt diese Anforderungen nur unvollständig. Es wertet alle Zitationen aus den im Index enthaltenen Zeitschriften aus, d. h. auch solche, die auf Buchpublikationen verweisen. Monographien und Sammelbände werden erfasst, sofern sie in den jeweiligen Journals zitiert wurden. Ein Problem liegt aber in der mangelhaften Abdeckung im Bereich der deutschsprachigen Zeitschriften, was zu systematischen Verzerrungen führt (vgl. Clermont/Schmitz 2008). Dilger (2000) forderte daher auch aus diesem Grund die Entwicklung eines deutschsprachigen Zitationsindex, der aber bis heute nicht existent oder konkret in Planung ist.

In Ermangelung eines speziell für die Zwecke der deutschsprachigen Sozialwissenschaften zugeschnittenen Zitationsindex kann *Google Scholar* als am ehesten geeignete Alternative gelten. Die Datenbank ist leistungsstark und kostenlos und wurde schon für eine Reihe von Studien erfolgreich eingesetzt (vgl. *Henrekson/Waldenström 2008; Dilger/Müller 2012; Müller 2010; Breuer 2009*). Aufgrund der breiten Abdeckung wissenschaftlicher Publikationen werden Zitationen von allen in alle Medientypen abgebildet. Auch Buch-zu-Buch-Verweise werden erfasst, die beim *ISI Web of Science* konstruktionsbedingt unberücksichtigt bleiben. Weil Veröffentlichungsdatenbanken von Wissenschaftsverlagen und Daten aus dem Internet automatisch durchsucht und ausgewertet werden, enthält *Google Scholar* in Bezug auf die BWL mehr Publikationen als die direkten Konkurrenten *ISI Web of Science* oder *Scopus* (vgl. *Harzing/van der Wal 2008*). Das Abfragen der Informationen kann mittels des von Anne-Wil Harzing entwickelten und kostenlos verfügbaren Programms *Publish or Perish* vorgenommen werden.⁷ Es bietet eine benutzerfreundliche Bedienoberfläche und errechnet zahlreiche bibliometrische Kennzahlen.

Abbildung 1: Zitationsanalyse mit Publish or Perish

The screenshot displays the 'Harzing's Publish or Perish' software interface. The main window is titled 'Multi-query center' and contains a table of queries. The 'Results' section shows the following summary statistics:

Papers:	9	Cites/paper:	3.56	h-index:	3	AWCRI:	6.25
Citations:	32	Cites/author:	9.58	g-index:	5	AW-index:	2.50
Years:	6	Papers/author:	2.37	hI-index:	3	AWCRIqA:	1.87
Cites/year:	5.33	Authors/paper:	3.89	hI, norm.:	2	e-index:	4.36
						hni-index:	1.17

The bottom table shows the following data:

Cites	Per year	Rank	Authors	Title	Year	Publication	Publi
132	216.00	1	R. Wodak, M Meyer	Methods for Critical Discourse Analysis	2009		book
20	6.67	2	T Fleiner-Gerster, M...	New Developments in Humanitarian ...	2008	... and Comparative Law ...	Cont.
16	3.20	3	M Aguilar, A Bhattach...	The new global challengers: how 10...	2006	Boston Consulting Group	
12	2.00	4	H Kasper, M Meyer, ...	Managers dealing with work-family-c...	2005	Journal of Managerial ...	emer
9	1.50	6	W Mayrhofer, J Ske...	Graduates' career aspirations and in...	2005	Human Resource ...	inger
9	3.00	5	MA Meyer	Liability of Prisoners of War for Offe...	2008	International and Comparat...	Cont.
8	2.67	7	M Meyer	The Compact Bedford Introduction L...	2008		Bedf...

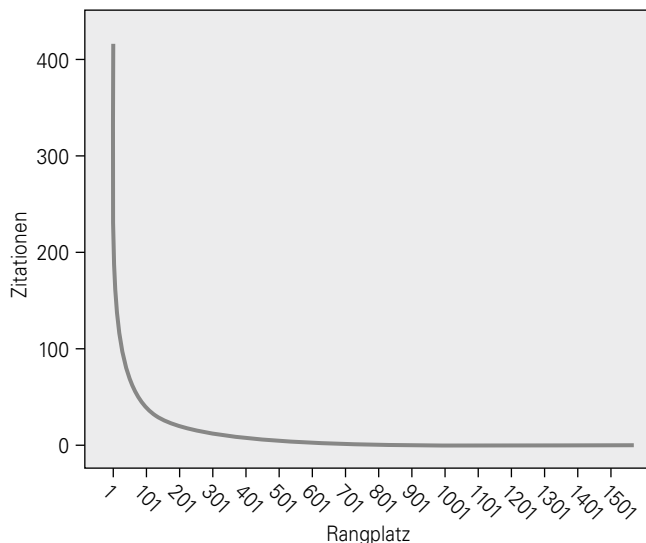
Schließlich stellt sich noch die Frage nach der Wertung von Publikationen mit mehreren Autoren. Hier ist die proportionale Aufteilung der Zitationen unter den Autoren einer Publikation ($1/n$) der am besten geeignete und daher etablierte Weg (vgl. *Dilger 2010*). Bleibt eine solche Korrektur um Koautorenschaften nämlich aus, so entsteht ein starker Anreiz zum Poolen von Veröffentlichungen, d. h. zwei Wissenschaftler gewinnen dadurch, dass sie einander gegenseitig mit als Koautoren aufnehmen. Dieses Problem besteht nicht nur bei zitationsbasierten Rankings, sondern auch bei

⁷ Kostenloser Download unter www.harzing.com/pop.htm (letzter Abruf am 27.08.2011).

der Verwendung anderer Kriterien wie etwa im Falle des Handelsblatt-Rankings (vgl. Müller 2010). Die proportionale Berücksichtigung von Koautorenschaften kann in der Auswertungssoftware *Publish or Perish* problemlos umgesetzt werden.⁸

Die zitationsbasierte Methode zum Erstellen von Forschungsleistungsrankings ist international bereits etabliert. Dabei sei beispielhaft auf das umfassende Ranking von Podsakoff et al. (2008) für die US-amerikanische BWL verwiesen. Für die deutschsprachige BWL wurde von Dilger/Müller (2012) ein Ranking auf der Grundlage von *Google Scholar* vorgelegt. Hierfür wurden die Publikationen der Jahre 2005 bis 2009 sämtlicher Mitglieder des Vereins der Hochschullehrer für Betriebswirtschaftslehre (VHB) berücksichtigt. Die darauf entfallenden Zitationen wurden mit Hilfe der Datenbank *Google Scholar* erhoben und um Koautorenschaften bereinigt (1/n). Sowohl bei Podsakoff et al. (2008) wie auch bei Dilger/Müller (2012) zeigt sich dabei eine extreme Ungleichverteilung der Zitationenwerte (siehe Abb. 2). Etwa 20 Prozent der Wissenschaftler akkumulieren mehr als 80 Prozent der Zitationen, während die übergroße Mehrheit der Forscher im Mittelfeld rangiert. Diese sogenannte Pareto-Verteilung (benannt nach Pareto 1896; siehe auch Hardy 2010) führt dazu, dass nur in den vorderen Rangplätzen signifikante Zitationsdifferenzen zwischen den einzelnen Rängen auftreten, während diese Unterschiede entlang der Verteilung immer unbedeutender werden.

Abbildung 2: Pareto-Verteilung der Zitationen



Quelle: Dilger/Müller 2012

⁸ Eine weitere Differenzierung hinsichtlich der Reihenfolge der Autoren erscheint hingegen nur dann sinnvoll, wenn davon ausgegangen werden kann, dass diese nach ihrem Arbeitsanteil an der Veröffentlichung und nicht alphabetisch sortiert sind. In der BWL ist dies jedoch nicht gegeben.

Insgesamt waren für etwas mehr als die Hälfte der 1.572 VHB-Mitglieder mehr als null Zitationen in *Google Scholar* nachweisbar, wobei aber berücksichtigt werden muss, dass im Datensatz auch viele bereits pensionierte Forscher enthalten sind. Aufgrund der Pareto-Verteilung lassen sich die Platzierungen jedoch nur im linken Bereich der Verteilung sinnvoll interpretieren, während im rechten Bereich die Unterschiede insignifikant sind. Die Tabelle zeigt die Zitationswerte für die fünf erstplatzierten Forscher. Zwischen dem ersten und dem fünften Platz liegen etwas mehr als 200 Zitationen, ebenso groß ist die Differenz zwischen Platz fünf und dem letzten Platz. Vergleichbare Informationen zur Verteilung liegen für das zeitschriftenbasierte Handelsblatt-BWL-Ranking nicht vor. Es ist jedoch anzunehmen, dass auch diese Verteilung der 80/20-Regel folgt, weil nicht erst die Zitationen, sondern typischerweise bereits die Publikationen höchst ungleich verteilt sind (vgl. *Podsakoff u. a. 2008*). Die Tatsache, dass eine Minderheit der Autoren für den Großteil der Publikationen verantwortlich ist, wurde bereits von Lotka (1926) beobachtet und wird deshalb auch als Lotkas Gesetz bezeichnet. Es dürfte sich in allen Formen von Forschungsleistungsrankings niederschlagen und dafür sorgen, dass die Unterschiede zwischen den ersten Rängen recht groß sind, während die große Gruppe der Forscher in einem nur schwer untereinander abgrenzbaren Mittelfeld rangiert.

Tabelle: Top 5 des zitationsbasierten Rankings von Dilger/Müller (2012)

R	Name	Hochschule	Publ.	Zit./Aut.
1	Christian Leuz	University of Chicago	33	415,43
2	Hartmut Stadler	Universität Hamburg	14	240,00
3	Martin Weber	Universität Mannheim	47	235,85
4	Christian Homburg	Universität Mannheim	38	211,34
5	Erik E. Lehmann	Universität Augsburg	34	209,66

Gegen zitationsbasierte Rankings werden in der Literatur einige Punkte vorgebracht: Ein im Rahmen der Zitationsanalyse häufig zu beobachtendes Phänomen ist der Matthäus-Effekt (siehe Merton 1968).⁹ Er beschreibt die Tatsache, dass bereits erfolgreiche Forscher besonders häufig zitiert werden, da sich die Zitierenden gern auf eine Autorität des Fachs beziehen. Daneben scheint es im Zitierverhalten eine Präferenz für methodologische Publikationen und Überblickswerke zu geben (vgl. *Voeth/Gawantka/Chatzopoulou 2006*). Beide Einwände richten sich bei näherer Betrachtung aber nicht gegen das zitationsbasierte Ranking im Speziellen, sondern berühren eine grundlegende Frage der Mess- und Vergleichbarkeit wissenschaftlicher Leistungen. Auch bei zeitschriftenbasierten Rankings würden diese Phänomene nämlich bereits etablierte und methodologisch ausgerichtete oder Überblicksartikel enthaltende Zeitschrif-

⁹ Benannt nach Mt. 25, 29: „Denn wer da hat, dem wird gegeben werden, dass er in Fülle habe“.

ten begünstigen. Das gilt insbesondere dann, wenn die Zeitschriftenbewertung auf *Impact*-Faktoren basiert. Ähnliches kann aber auch für befragungsbasierte Zeitschriftenrankings plausibel angenommen werden.

Da Rankings naturgemäß vergangenheitsbezogen sind, gibt es eine Präferenz für ältere Veröffentlichungen, weil diese bereits auf eine längere Zitationsgeschichte zurückblicken. Das ist bei der Wahl eines geeigneten Untersuchungszeitraums zu berücksichtigen, sollte jedoch nicht voreilig als Fehler oder Verzerrung der Methodik interpretiert werden. Es ist vielmehr ein grundlegendes Charakteristikum von Wissenschaft, dass Ideen und Erkenntnisse erst im Nachhinein in ihrer wahren Tragweite abgeschätzt werden können.¹⁰ Eine wissenschaftliche Veröffentlichung geht vielmehr durch einen Lebenszyklus, zu dessen Beginn die Zitationen stark zunehmen, einen Höhepunkt erreichen und dann langsam wieder fallen, weil die veröffentlichten Inhalte entweder keine Relevanz mehr haben oder in das Allgemeinwissen des Fachs übergegangen sind und deswegen nicht mehr zitiert werden müssen (vgl. *Aizenman/Kletzer 2008*).

Ein praktischer Nachteil zitationsbasierter Rankings ist der im Vergleich zu zeitschriftenbasierten Rankings ungleich größere Aufwand, den ihre Erstellung erfordert (vgl. *Dilger 2010*): Erstens müssen für jede Publikation Zitationskennzahlen aus der Datenbank entnommen werden. Zweitens müssen sämtliche Informationen für jede Neuauflage des Rankings komplett neu erhoben werden, da auch alle älteren Veröffentlichungen weiter zitiert werden können. Automatisierte Software-Tools sind gegenwärtig weder für das ISI *Web of Science* noch für *Google Scholar* verfügbar.

Ein gravierender Nachteil zitationsbasierter Rankings liegt in ihrer Anreizstruktur im Hinblick auf mögliches strategisches Verhalten. Selbstzitationen und Zitierkartelle könnten von Wissenschaftlern dazu genutzt werden, die eigene Rankingposition zu verbessern (vgl. *Dilger 2000*). Die Datenbanken können zwar Selbstzitationen grundsätzlich filtern. In *Google Scholar* geschieht dies aber – trotz anderweitiger Behauptungen von *Google* –, wenn überhaupt, nur sehr unzureichend (vgl. *Henrekson/Waldenström 2008*). Zitierkartelle sind hingegen technisch nicht zu erkennen und im Übrigen von korrektem wissenschaftlichen Verhalten auch kaum trennscharf abzugrenzen: Wenn sich nur sehr wenige Forscher mit einem eng umgrenzten Thema beschäftigen und sich in ihren Publikationen aufeinander beziehen, deutet dies nicht zwangsläufig auf ein Zitierkartell hin (vgl. *Dyckhoff/Thieme/Schmitz 2005*). Zum gegenwärtigen Zeitpunkt dürften die Anreizwirkungen zitationsbasierter Rankings ohnehin noch keine große Rolle spielen, da sie in den deutschsprachigen Sozialwissenschaften sehr ungebräuchlich sind. Erwartet werden muss strategisches Verhalten erst dann,

¹⁰ So wird der sogenannte Wirtschaftsnobelpreis typischerweise für Erkenntnisse verliehen, die bereits seit Jahrzehnten publiziert sind.

wenn ein zitationsbasiertes Ranking große Aufmerksamkeit findet und das Abschneiden darin Konsequenzen für die Gerankten hat.

3.2.3 Weitere Methoden

Neben den beiden hier vorgestellten Verfahren sind weitere Kriterien denkbar, nach denen Personen oder Hochschulen in eine Reihenfolge gebracht werden können. Für ein Ranking von Fakultäten oder Fachbereichen ließe sich z. B. auch der befragungs-basierte Ansatz analog zum VHB-Jourqual Zeitschriftenranking anwenden, bei dem Hochschullehrer direkt nach der wahrgenommenen Reputation der untersuchten Institutionen befragt würden. Ein solches Vorgehen wäre aber aus mehreren Gründen wenig erfolgversprechend und dürfte deshalb noch nicht systematisch angewandt worden sein: So bestünde ein Anreiz zu strategischen Antworten zugunsten der eigenen Hochschule; große Institutionen könnten allein aus ihrer Größe einen Bekanntheitsvorteil ziehen. Schließlich ist das Verfahren mit einem nicht zu unterschätzenden Durchführungsaufwand verbunden, so dass auch daher eine Anwendung auf Individual-ebene nicht möglich erscheint.

CHE-Forschungsranking als multidimensionales Verfahren

Im Rahmen des multidimensionalen Forschungsrankings des Centrums für Hochschul-entwicklung (CHE) wurde die Befragungsmethode als eine von mehreren Dimensionen gewählt (*vgl. hierzu im Folgenden Berghoff u. a. 2009*). BWL-Professoren wurden dazu befragt, wie sie die Reputation der anderen Hochschulen in der Forschung einschätzen. Als weitere Indikatoren für Forschungsleistung wurde u. a. die Anzahl an Promotionen sowie die Höhe und Struktur der verausgabten Drittmittel herangezogen. In Bezug auf die Promotionen kann hinterfragt werden, ob die Betreuung von Doktoranden nicht eher eine Form von Lehrleistung als ein Forschungsindikator ist. Zudem wird die Qualität der Promotionen nicht berücksichtigt. Hinsichtlich der Drittmittel ist unklar, ob es sich überhaupt um ein Leistungsmaß handelt. Versteht man wissenschaftliche Forschung als Produktionsprozess von Wissen, so sind sie als Inputgröße anzusehen und sollten bei einer Outputmessung unberücksichtigt bleiben (*vgl. Rassenhövel/Dyckhoff 2006*).¹¹

¹¹ Dagegen kann eingewandt werden, dass die Gewährung von Drittmitteln (z. B. bei der DFG) nach einem Antragsverfahren erfolgt, in dem die bereits erbrachten Forschungsleistungen eine wichtige Rolle spielen. Auf diese Weise würden sie zu einem mittelbaren Qualitätsindikator. Ob jedoch die Qualität von Drittmittelanträgen in jedem Fall mit der Qualität von Forschungsleistungen gleichgesetzt werden kann, erscheint fraglich, weil die Beteiligung an komplexen Antragsverfahren für die Forscher sehr zeitintensiv ist und damit zulasten der für die eigentliche Forschung aufgewandten Zeit gehen kann (*vgl. hierzu Rassenhövel/Dyckhoff 2006*). Daneben müssen innovative Forschungsprojekte zumindest in den Sozial- und Geisteswissenschaften nicht notwendigerweise viele Ressourcen verbrauchen, d. h. die Einwerbung von Drittmitteln ist nicht in jedem Fall gleich attraktiv.

Weitere Indikatoren des CHE-Rankings waren in der Betriebs- und Volkswirtschaftslehre deutschsprachige und internationale Publikationen. Im Rahmen der Publikationsanalyse wurde ein gemischtes Verfahren aus zeitschriftenbasierter Qualitätsgewichtung und reiner Zählung der Beiträge angewandt. Vor dem Hintergrund der eingangs erläuterten Überlegungen kann ein solches Vorgehen durchaus hinterfragt werden. So ist z. B. nicht einsichtig, warum internationale und nationale Publikationen zwei getrennte Indikatoren darstellen sollen. Inhaltlich unklar bleibt ferner, wieso erstere lediglich gezählt, letztere aber auf Koautorenschaften und Seitenzahl normiert und anhand des VHB-Jourqual qualitätsgewichtet werden.

Alle Indikatoren des CHE-Rankings werden schließlich in einem *Scoring*-Modell aggregiert, das die Forschungsleistung der betriebswirtschaftlichen Fachbereiche multidimensional abbilden soll. Dabei fällt auf, dass die Korrelationen zwischen den unterschiedlichen Indikatoren vielfach nicht signifikant und auch sonst meist nur sehr schwach ausgeprägt sind (vgl. *Berghoff u. a. 2009*). Vor dem Hintergrund dieser Ergebnisse scheint ein gewichtetes Aufaddieren der in verschiedene Richtungen deutenden Indikatoren nur bedingt sinnvoll: Aufgrund der Divergenz der Zwischenergebnisse ist zu vermuten, dass das resultierende Ranking stark von der Gewichtung im *Scoring*-Modell abhängt. In jedem Fall sind die Ergebnisse aufgrund der Vielzahl der verwendeten und miteinander verknüpften Methoden kaum noch transparent zu machen. Das hier nur in seinen grundsätzlichen Zügen wiedergegebene Verfahren verdeutlicht, dass eine Vielzahl möglicher Indikatoren zur Messung von Forschungsleistung existiert. Daneben sind noch viele weitere Kriterien vorstellbar. Vor diesem Hintergrund scheint eine begründete Auswahl sinnvoller als der Versuch, sämtliche möglichen Verfahren in ein Ranking zu implementieren.

Die anhand des CHE-Rankings vorgestellten Alternativindikatoren zur Messung von Forschungsleistung ließen sich grundsätzlich nicht nur auf Hochschulen, sondern auch auf Lehrstühle, Institute oder auf Personen anwenden. Anhand von Drittmittelverwendung, Anzahl der Promotionen, mittels einer Befragung (oder analog zum CHE-Modell durch eine Kombination aller genannten Verfahren) ließen sich Personenrankings erstellen. Ob ein solches Vorgehen zielführend ist, kann mit Blick auf die Ausführungen zum CHE-Hochschulranking allerdings stark bezweifelt werden.

Herausgebertätigkeit als Qualitätsindikator

Einen anderen Ansatz wählen *Frey/Rost (2010)*, indem sie wissenschaftliche Qualität durch die Herausgebertätigkeit eines Forschers für hochrangige wissenschaftliche Zeitschriften (die im *ISI Web of Knowledge Journal Citation Report* oder im VHB-Jourqual Ranking mit AA+ bis B bewertet wurden) abbilden. Wenngleich man den Autoren darin zustimmen muss, dass auch die Administration und Redaktion von wissenschaftlichen Zeitschriften eine wichtige Aufgabe eines Hochschullehrers ist, so bleibt letztlich

offen, wieso diese Methodik den an den Publikationen der Forscher orientierten Verfahren überlegen sein soll. Betrachtet man die Produktion von Wissen als primären Prozess einer Hochschule (neben der Lehre) oder einer wissenschaftlichen Forschungseinrichtung, so käme der Organisation des Publikationsverfahrens eher die Rolle eines Sekundär- oder Hilfsprozesses zu (vgl. Porter 1985). Zwar soll nicht bestritten werden, dass sich die qualifiziertesten Forscher eines Fachs besonders für die Tätigkeit als Herausgeber einer Zeitschrift empfehlen. Dennoch sollte sich die Messung von Forschungsleistung so eng wie möglich am Inhalt bzw. den wissenschaftlichen Erkenntnissen an sich und weniger an den dazu notwendigen Hilfsprozessen orientieren.

Von diesen eher konzeptionellen Überlegungen abgesehen hat das von Frey/Rost entworfene Verfahren praktisch wenig Aussagekraft. So gibt es in einem weltweiten Ranking der führenden Ökonomen nur sieben (sic!) mögliche Rangplätze. Angewandt auf die deutschsprachige BWL ließe sich vermuten, dass sich nahezu alle Forscher auf einem geteilten siebten Platz wiederfänden. Der Ansatz scheint also in Bezug auf die Messung von Forschungsleistungen einer Vielzahl von Forschern ungeeignet. Eine mögliche Erweiterung der untersuchten Herausgeberkreise könnte zwar für mehr Forscher zu Ergebnissen führen, wirft aber die Frage auf, wie die Herausgeberschaften in unterschiedlichen Zeitschriften gegeneinander gewichtet werden sollen. So wäre die Tätigkeit als Herausgeber von beispielsweise WISU sicherlich weniger reputierlich als die für das *Journal of Economic Literature*. Diese Frage wird von Frey/Rost nicht gelöst, sondern durch die Begrenzung auf wenige internationale Top-Journals umgangen.

4 Fazit

Die vergleichende Analyse der methodologischen Ansätze zur Erstellung von Forschungsleistungsrankings hat gezeigt, dass die Betrachtung der Zitationen einer Veröffentlichung gegenüber den anderen besprochenen Methoden zwei wesentliche konzeptionelle Vorteile aufweist: Erstens ist dieses Verfahren nicht notwendigerweise auf Zeitschriftenpublikationen begrenzt, sondern es können auch Monographien und Sammelbände betrachtet werden, und zweitens können die verfügbaren Daten bestmöglich ausgenutzt werden, weil direkt an der einzelnen Publikation angeknüpft wird. Diese konzeptionellen Vorteile dürften die praktischen Nachteile, die in den Anforderungen an die Datenbank und im Erstellungsaufwand liegen, langfristig mehr als aufwiegen.

Bei der Interpretation der Ergebnisse besonders kritisch ist die Frage, ob einzelne Rangplätze oder geringfügige Differenzen in den als Rankingkriterien verwendeten Kennzahlen als signifikante Leistungsunterschiede bewertet werden können. Dieses Problem stellt sich unabhängig davon, ob Zitationen oder auf anderer Grundlage er-

rechnete Punktwerte (wie z. B. beim Handelsblatt-Ranking) zur Rangfolgenbildung herangezogen werden. Die im Ranking (notwendigerweise) vergebenen Rangplätze und die bis auf ihre Nachkommastellen berechneten Kennzahlen suggerieren dem nicht mit der Methodologie vertrauten Leser eine Scheingenaugigkeit, die mit Blick auf die verwendeten Methoden und die für die Bibliometrie typischen Verteilungsgesetze nicht gerechtfertigt scheint. Daher ist die Interpretation geringfügiger Unterschiede hinsichtlich der Zitationsindikatoren oder Rangplätze auch wenig sinnvoll (siehe Abschnitt 3.2.2). Trotz dieser Unschärfe im Detail lassen sich aber sehr wohl Aussagen über wesentliche Differenzen machen. Vor allem die Unterschiede zwischen den bestplatzierten Forschern und der breiten Masse deuten auf real existierende Leistungsunterschiede hin.

Im Rahmen dieser Untersuchung wurde vornehmlich aus der Perspektive der BWL argumentiert. Die Kontroverse über die Outputmessung an Forschungsinstitutionen ist aber keineswegs auf dieses Fach begrenzt, sondern steht in mehr oder weniger starker Ausprägung auch in anderen Sozialwissenschaften zur Debatte (*Hofmeister/ Ursprung 2008*; siehe für die Soziologie z. B. *Litzenberger/Sternberg 2005*; *Gläser 2006*; für die Politikwissenschaft beispielsweise *Plümper 2003*; *Schneider 2007*; *Erne 2007*; exemplarisch für die Volkswirtschaftslehre *Ursprung/Zimmer 2007*; *Graber/Launov/ Walde 2008*). Daher stellt sich die Frage, inwieweit sich die hier am Beispiel der BWL vorgenommenen Überlegungen auch auf die Nachbardisziplinen ausweiten lassen. Mit Blick auf die zentralen Charakteristika der Forschungskommunikation in den genannten Fächern spricht einiges dafür, dass zitationsbasierte Methoden zur Forschungsleistungsmessung auch in diesen Fächern sinnvoll angewandt werden können. So sind Monographien und andere Buchpublikationen auch in diesen Disziplinen neben den wissenschaftlichen Zeitschriften bedeutende Medien (*vgl. Thompson 2002*), deren Nichtbeachtung bei einem ausschließlich auf Zeitschriftenpublikationen basierenden Bewertungsverfahren ebenso problematisch wäre. Zweitens haben auch in diesen Wissenschaften deutschsprachige Veröffentlichungen nach wie vor Gewicht (*vgl. Plümper/Schimmelpfennig 2007*), wenngleich dies themenspezifisch unterschiedlich oder auch abnehmend sein mag. Die Abdeckungsbreite traditioneller, nahezu ausschließlich auf die englischsprachige Literatur fokussierter Literatur- und Zitationsdatenbanken dürfte auch hier zu Problemen führen (*vgl. Hicks 1999*), sodass *Google Scholar* bessere Ergebnisse verspricht, zumal bereits bei den englischsprachigen Publikationen im Bereich der Sozialwissenschaften eine deutlich größere Abdeckung zu verzeichnen ist (*vgl. Harzing 2010*). Die Argumente hinsichtlich der Herausgeber-tätigkeit für Fachzeitschriften und der Drittmittel als alternative Rankingkriterien dürften ebenso auch in den Nachbardisziplinen der BWL gelten. Vor diesem Hintergrund scheinen Zitationen auch hier als Maß für den Einfluss und die Qualität wissenschaftlicher Publikationen geeignet zu sein.

Die Eignung einer Methode muss jedoch immer vor dem Hintergrund der Ziele des jeweils aufzustellenden Rankings beurteilt werden. Soll beispielsweise ein Bewerber-ranking als Hilfsmittel zur (Vor-)Auswahl im Rahmen eines Berufungsverfahrens erstellt werden, so erscheint der zitationsbasierte Ansatz als der am ehesten geeignete. Der im Vergleich zum zeitschriftenbasierten Ranking höhere Aufwand dürfte bei einer begrenzten Zahl von Personen nicht ins Gewicht fallen und auch die Eignung der dazu heranzuziehenden Zitationsdatenbanken lässt sich anhand der Abdeckungsbreite leicht beurteilen, wenn man die Suchergebnisse mit den Publikationslisten der Bewerber abgleicht. Fehler in der Datenbank, wie sie insbesondere für *Google Scholar* dokumentiert sind (vgl. *Jacsó 2006*), lassen sich unmittelbar erkennen und oft bereits mit der Auswertungssoftware *Publish or Perish* manuell korrigieren. Dennoch mag der zitationsbasierte Ansatz nicht in jedem Fall die Methode der Wahl sein. Die auf Zeitschriftengewichten basierende Methode zeichnet sich nämlich durch eine ex ante größere Planbarkeit aus, wie sie z. B. bei Vereinbarungen über Leistungszulagen im Rahmen der W-Besoldung gegeben sein sollte. Wenn eine Hochschulleitung derartige Instrumente trotz der mit ihnen verbundenen Probleme in der Professorenbesoldung nutzen möchte, müssen die vereinbarten Ziele von den Forschern sinnvoll angestrebt werden können, was bei einem Artikel in einer bestimmten Zeitschriftenkategorie im Gegensatz zu einer definierten Zahl von Zitationen eher möglich erscheint.

Da die anzuwendenden Methoden vor dem Hintergrund unterschiedlicher Ziele bewertet werden sollen, folgt daraus ein Nebeneinander verschiedener Rankings. Keine Methodik dominiert bislang sämtliche Alternativen in allen Eigenschaften und folglich dürfen die Ergebnisse immer nur innerhalb der methodeninduzierten Grenzen interpretiert werden. Wenn methodisch unterschiedliche Rankings mit voneinander abweichenden Ergebnissen nebeneinander existieren, kann dies jedoch für den Fortschritt der Bibliometrie durchaus hilfreich sein, denn eine weitere Diskussion und die Weiterentwicklung der Methoden sollten sich auf die Validität der Messinstrumente positiv auswirken. Nicht zuletzt ist ein Nebeneinander mehrerer gleichberechtigter Ansätze auch dazu geeignet, die von Kritikern befürchteten Fehlanreize in gewissen Grenzen zu halten. Die befürchtete Zielverschiebung der Gerankten weg von wissenschaftlicher Exzellenz hin zu einer Optimierungsstrategie für ein bestimmtes Ranking ist aus der Sicht der Betroffenen weniger attraktiv, wenn es mehrere Rankings mit verschiedenen Metriken gibt. In diesem Sinne ist die Entwicklung alternativer und innovativer bibliometrischer Methoden weiterhin sehr begrüßenswert.

Literatur

Adler, Nancy J.; Harzing, Anne-Wil (2009): When knowledge wins: Transcending the sense and nonsense of academic rankings. In: *The Academy of Management Learning and Education* 8, 2009, S. 72–95.

Aizenman, Joshua; Kletzer, Kenneth (2008): The Life Cycle of Scholars and Papers in Economics: the 'Citation Death Tax'. NBER Working Paper No. 13891, Cambridge (MA).

Albers, Sönke (2009): Misleading rankings of research in business. In: *German Economic Review* 10, 2009, S. 352–363.

Berghoff, Sonja; Federkeil, Gero; Giebisch, Petra; Hachmeister, Cort-Denis; Hennings, Mareike; Roessler, Isabel; Ziegele, Frank (2009): Das CHE-Forschungsranking deutscher Universitäten 2009. Gütersloh.

Blomeyer, Christian (2007): Professorenbesoldung. Hamburg.

Breuer, Wolfgang (2009): Google Scholar as a Means for Quantitative Evaluation of German Research Output in Business Administration: Some Preliminary Results. Diskussionspapier, Aachen, im Internet unter <http://ssrn.com/abstract=1280033> (letzter Aufruf am 25.11.2011).

Campbell, Frank (1896): The Theory of National and International Bibliography. London.

Clermont, Marcel; Schmitz, Christian (2008): Erfassung betriebswirtschaftlich relevanter Zeitschriften in den ISI-Datenbanken sowie der Scopus-Datenbank. In: *Zeitschrift für Betriebswirtschaft* 78, 2008, S. 987–1010.

Cole, Francis Joseph; Eales, Nellie B. (1917): A History of Comparative Anatomy: Pt. 1. A Statistical Analysis of the Literature. In: *Science Progress* 11, 1917, S. 578–596.

Cronin, Blaise (1982): Norms and functions in citation: The view of journal editors and referees in psychology. In: *Social Science Information Studies* 2, 1982, S. 65–77.

Dilger, Alexander (2010): Rankings von Zeitschriften und Personen in der BWL. In: *Zeitschrift für Management* 5, 2010, S. 91–102.

Dilger, Alexander (2000): Plädoyer für einen Sozialwissenschaftlichen Zitationsindex. In: *Die Betriebswirtschaft* 60, 2000, S. 473–484.

Dilger, Alexander; Müller, Harry (2012): Ein Forschungsleistungsranking auf der Grundlage von Google Scholar. Wird erscheinen in: *Zeitschrift für Betriebswirtschaft*.

Dyckhoff, Harald; Rassenhövel, Silvia; Gilles, Roland; Schmitz, Christian (2005): Beurteilung der Forschungsleistung und das CHE-Forschungsranking betriebswirtschaftlicher Fachbereiche. In: *Wirtschaftswissenschaftliches Studium* 34, 2005, S. 62–69.

Dyckhoff, Harald; Schmitz, Christian (2007): Forschungsleistungsmessung mittels SSCI oder SCI-X? Internationale Sichtbarkeit und Wahrnehmung der Betriebswirtschaftslehre von 1990 bis 2004. In: *Die Betriebswirtschaft* 67, 2007, S. 638–662.

Dyckhoff, Harald; Thieme, Annegret; Schmitz, Christian (2005): Die Wahrnehmung deutschsprachiger Hochschullehrer für Betriebswirtschaft in der internationalen Forschung: Eine Pilotstudie zu Zitationsverhalten und möglichen Einflussfaktoren. In: *Die Betriebswirtschaft* 65, 2005, S. 350–372.

Erne, Roland (2007): On the use and abuse of bibliometric performance indicators: a critique of Hix's 'global ranking of political science departments'. In: *European Political Science* 6, 2007, S. 306–314.

Frey, Bruno S. (2003): Publishing as prostitution? Choosing between one's own ideas and academic success. In: *Public Choice* 116, 2003, S. 205–223.

Frey, Bruno S.; Rost, Katja (2010): Do rankings reflect research quality? In: *Journal of Applied Economics* 13, 2010, S. 1–38.

Garfield, Eugene (1979): Is citation analysis a legitimate evaluation tool? In: *Scientometrics* 1, 1979, S. 359–375.

Gläser, Jochen (2006): Die Fallstricke der Bibliometrie. In: *Soziologie* 35, 2006, S. 42–51.

Graber, Michael; Launov, Andrey; Walde, Klaus (2008): Publish or Perish? The Increasing Importance of Publications for Prospective Economics Professors in Austria, Germany and Switzerland. In: *German Economic Review* 9, 2008, S. 457–472.

Griffith, David A.; Tamer Cavusgil, S.; Xu, Shichun (2008): Emerging themes in international business research. In: *Journal of International Business Studies* 39, 2008, S. 1220–1235.

Hardy, Michael (2010): Pareto's Law. In: *Mathematical Intelligencer* 32, 2010, S. 38–43.

Harzing, A. W. (2010): The publish or perish book: your guide to effective and responsible citation analysis. Melbourne.

Harzing, Anne-Wil; van der Wal, Ron (2008): Google Scholar as a new source for citation analysis. In: *Ethics in Science and Environmental Politics* 8, 2008, S. 61–73.

Hennig-Thurau, Thorsten; Walsh, Gianfranco; Schrader, Ulf (2004): VHB-JOURQUAL: Ein Ranking von betriebswirtschaftlich-relevanten Zeitschriften auf der Grundlage von Expertenurteilen. In: *Zeitschrift für betriebswirtschaftliche Forschung* 56, 2004, S. 520–545.

Henrekson, Magnus; Waldenström, Daniel (2008): How Should Research Performance be Measured? Evidence from Rankings of Academic Economists. In: *Working Paper Series in Economics and Finance* 2008.

Hicks, D. (1999): The difficulty of achieving full coverage of international social science literature and the bibliometric consequences. In: *Scientometrics* 44, 1999, S. 193–215.

Hofmeister, Robert; Ursprung, Heinrich W. (2008): Das Handelsblatt Ökonomen-Ranking 2007: Eine kritische Beurteilung. In: *Perspektiven der Wirtschaftspolitik* 9, 2008, S. 254–266.

Jacsó, Peter (2006): Deflated, inflated and phantom citation counts. In: *Online information review* 30, 2006, S. 297–309.

Kieser, Alfred (2010): Unternehmen Wissenschaft? In: *Leviathan* 38, 2010, S. 347–367.

Klingemann, Hans-Dieter (1988): Zitierhäufigkeit als Qualitätsindikator: Eine Rangordnung der amerikanischen politikwissenschaftlichen Fachbereiche in den 80er Jahren. In: Daniel, H. D.; Fisch, R. (Hrsg.): *Evaluation von Forschung*. Konstanz, S. 201–214.

Lawrence, Peter A. (2003): The politics of publication. In: *Nature* 422, 2003, S. 259–261.

Le Pair, Cees (1988): The citation gap of applicable science. In: van Raan, Anthony F. J. (Hrsg.): *Handbook of Quantitative Studies of Science and Technology*. Amsterdam, S. 537–553.

Litzenberger, Timo; Sternberg, Rolf (2005): Die Forschungsleistung der Soziologie an zehn deutschen Universitäten. In: *Soziologie* 34, 2005, S. 174–190.

Lotka, A. J. (1926): The frequency distribution of scientific productivity. In: *Journal of the Washington Academy Sciences* 16, 1926, S. 317–323.

MacRoberts, Michael H.; MacRoberts, Barbara R. (1996): Problems of citation analysis. In: *Scientometrics* 36, 1996, S. 435–444.

Meadows, Arthur Jack (1974): *Communication in science*. London.

Merton, Robert K. (1968): The Matthew effect in science. In: *Science* 159, 1968, S. 56–63.

Müller, Anja; Storbeck, Olaf (2009): Ranking: Die BWL erfindet sich neu. In: *Handelsblatt* vom 20.5.2009.

Müller, Harry (2010): Wie valide ist das Handelsblatt-BWL-Ranking: Zeitschriften und zitationsbasierte Personenrankings im Vergleich. In: *Betriebswirtschaftliche Forschung und Praxis* 62, 2010, S. 152–166.

Musgrave, Richard A.; Musgrave, Peggy B. (1976): *Public finance in theory and practice*. New York.

Neuhaus, Christoph; Daniel, Hans-Dieter (2008): Data sources for performing citation analysis: an overview. In: *Journal of Documentation* 64, 2008, S. 193–210.

Osareh, Farideh (1996): Bibliometrics, citation analysis and co-citation analysis: A review of literature I. In: *Libri* 46, 1996, S. 149–158.

Pareto, Vilfredo (1896): *Cours d'économie politique*. Lausanne.

Pellens, Bernhard; Füllbier, Rolf Uwe; Gassen, Joachim; Sellhorn, Thorsten (2008): *Internationale Rechnungslegung. IFRS 1 bis 8, IAS 1 bis 41, IFRIC-Interpretationen, Standardentwürfe*. Stuttgart.

Plümper, Thomas (2003): Publikationstätigkeit und Rezeptionserfolg der deutschen Politikwissenschaft in internationalen Fachzeitschriften, 1990–2002. Eine bibliometrische Analyse der Veröffentlichungsleistung deutscher politikwissenschaftlicher Fachbereiche und Institute. In: Politische Vierteljahresschrift 44, 2003, S. 529–544.

Plümper, Thomas; Schimmelpfennig, Frank (2007): Wer wird Prof - und wann? Berufungsdeterminanten in der deutschen Politikwissenschaft. In: Politische Vierteljahresschrift 48, 2007, S. 97–117.

Podsakoff, Philip M.; MacKenzie, Scott B.; Podsakoff, Nathan P.; Bachrach, Daniel G. (2008): Scholarly influence in the field of management: A bibliometric analysis of the determinants of university and author impact in the management literature in the past quarter century. In: Journal of Management 34, 2008, S. 641–720.

Popper, Karl (2005): Logik der Forschung. Tübingen.

Porter, Michael E. (1985): Competitive advantage: creating and sustaining superior performance. New York.

Power, Michael (1997): The audit society: rituals of verification. New York.

Rassenhövel, Sylvia; Dyckhoff, Harald (2006): Die Relevanz von Drittmittelindikatoren bei der Beurteilung der Forschungsleistung im Hochschulbereich. In: Zelewski, Stephan; Akca, Naciye (Hrsg.): Fortschritt in den Wirtschaftswissenschaften: Wissenschaftstheoretische Grundlagen und exemplarische Anwendungen. Wiesbaden, S. 85–112.

Schlinghoff, Axel; Backes-Gellner, Uschi (2002): Publikationsindikatoren und die Stabilität wirtschaftswissenschaftlicher Zeitschriftenrankings. In: Zeitschrift für betriebswirtschaftliche Forschung 54, 2002, S. 343–362.

Schmitz, Christian (2008): Messung der Forschungsleistung in der Betriebswirtschaftslehre auf Basis der ISI-Zitationsindizes: Eine kritische Analyse anhand konzeptioneller Überlegungen und empirischer Befunde. Lohmar/Köln.

Schneider, Gerald (2007): Why is European Political Science so Unproductive and what should be done about It: a Symposium. In: European Political Science 6, 2007, S. 156–159.

Schulze, Günther G.; Warning, Susanne; Wiermann, Christian (2008): Zeitschriftenrankings für die Wirtschaftswissenschaften? Konstruktion eines umfassenden Metaindexes. In: Perspektiven der Wirtschaftspolitik 9, 2008, S. 286–305.

Seglen, Per O. (1997): Why the impact factor of journals should not be used for evaluating research. In: British Medical Journal 314, 1997, S. 497–502.

Seglen, Per O. (1994): Causal relationship between article citedness and journal impact. In: Journal of the American Society for Information Science 45, 1994, S. 1–11.

Süß, Stefan (2007): Grenzen einer leistungsorientierten Anreizgestaltung für Hochschullehrer. In: Hochschulmanagement 2, 2007, S. 68–72.

Thompson, Jennifer Wolfe (2002): The death of the scholarly monograph in the humanities? Citation patterns in literary scholarship. In: *Libri* 52, 2002, S. 121–136.

Ursprung, Heinrich W.; Zimmer, Markus (2007): Who Is the 'Platz-Hirsch' of the German Economics Profession? A Citation Analysis. In: *Jahrbücher für Nationalökonomie und Statistik* 227, 2007, S. 187–208.

Voeth, Markus; Gawantka, Axel; Chatzopoulou, Georgia (2006): Impact auf die deutsche Marketingforschung. In: *Marketing Zeitschrift für Forschung und Praxis* 28, 2006, S. 7–20.

Weingart, Peter (2005): Impact of bibliometrics upon the science system: Inadvertent consequences? In: *Scientometrics* 62, 2005, S. 117–131.

Weingart, Peter; Winterhager, Matthias (1984): Die Vermessung der Forschung: Theorie und Praxis der Wissenschaftsindikatoren. Frankfurt a.M./New York.

Wolf, Joachim; Rohn, Anne; Macharzina, Klaus (2006): Forschungsleistung in der deutschsprachigen Betriebswirtschaftslehre: Konzeption und Befunde einer empirischen Untersuchung. Wiesbaden.

Anschrift des Verfassers:

Harry Müller, B.A., M.Sc.
Westfälische Wilhelms-Universität Münster
Institut für Organisationsökonomik
Scharnhorststr. 100
48151 Münster
E-Mail: harry.mueller@uni-muenster.de

Harry Müller ist wissenschaftlicher Mitarbeiter am Institut für Organisationsökonomik der Universität Münster.

Blossfeld, Hans-Peter; Roßbach, Hans-Günther; von Maurice, Jutta (Hrsg.): Education as a Lifelong Process. The German National Education Panel Study (NEPS). Zeitschrift für Erziehungswissenschaften: Special Issue 14 / 2011, ISBN: 978-3-531-17785-4, 330 Seiten

In diesem Sonderheft der Zeitschrift für Erziehungswissenschaften wird in 20 Beiträgen eine der größten und anspruchsvollsten Studien der empirischen Bildungsforschung in Deutschland vorgestellt: Das Nationale Bildungspanel bzw. National Panel of Education (NEPS). Unterstützt durch das Rahmenprogramm des BMBF zur Förderung der empirischen Bildungsforschung und unter Mitwirkung der hochkarätigsten Bildungsforscher Deutschlands werden hochwertige quantitative Informationen über Bildungsverläufe und -ergebnisse für den gesamten Lebensverlauf erhoben – von der frühkindlichen Phase bis hin zu Hochschul- und Erwachsenenbildung. Um die Bildungsgeschichte in Abhängigkeit von Ausgangsbedingungen und institutionellen Umgebungsfaktoren im Längsschnitt untersuchen zu können, wählte das multidisziplinäre Team einen Kohortenansatz und vielfältige methodische Zugänge von biographischen Interviews bis hin zu eigens entwickelten computergestützten Kompetenztestverfahren. Das reichhaltige Datenmaterial wird für die wissenschaftliche Nutzung in Scientific Use Files zur Verfügung gestellt.

Die 20 Beiträge von NEPS-Wissenschaftlern stellen zunächst die konzeptuellen und theoretischen Hintergründe vor, insbesondere die fünf thematische „Säulen“ des NEPS (Kompetenzentwicklung, Lernumwelten, Bildungsentscheidungen, Migrationshintergrund und Bildungsrenditen). Weitere Kapitel behandeln Fragen der Datenqualität, der Stichprobenwahl, von Methoden der Kompetenzmessung bei Erwachsenen sowie des Datenschutzes und der Datenweitergabe. Darüber hinaus werden die Hintergründe und Vorgehensweise auf den acht „Bildungsetappen“ des NEPS in eigenen Kapiteln beschrieben: Neugeborene und frühkindliche institutionelle Betreuung, Kindergarten und Übergang in die Grundschule, Grundschule und Übergang in die Sekundarstufe I, Sekundarstufe I und Übergang in die Sekundarstufe II, Sekundarstufe II und Übergang in den tertiären Bereich oder Arbeitsmarkt, berufliche Ausbildung und Arbeitsmarkt, Hochschule und Übergang in den Beruf sowie Bildung im Erwachsenenalter bzw. Lebenslanges Lernen. Der Band illustriert hervorragend das enorme Datenpotential des NEPS. Er bietet allen Interessierten eine exzellente Ausgangsbasis für die Formulierung theoretisch fundierter Forschungsfragen und weiterführender Analysen.

Würmseer, Grit: Auf dem Weg zu neuen Hochschultypen. Wiesbaden: Verlag für Sozialwissenschaften, 2010, ISBN 978-3-531-17771-7, 361 Seiten

Die Hochschulen in Deutschland haben in jüngerer Vergangenheit weitgreifende Veränderungen durchlaufen. Grit Würmseer geht in ihrer Dissertation der Frage nach, inwieweit diese Reformen ihre Identität als Organisationen beeinflussen. Der erste Teil des Buchs stellt einleitend die aktuellen Reformen im Hochschulbereich vor und gibt eine theoretische Fundierung von Organisationen und ihrer Identität auf Basis organisationssoziologischer Modelle. Der zweite Teil enthält die Ergebnisse einer eigenen qualitativen Untersuchung. Hierzu wurden Interviews mit mehreren Personen vorrangig aus der Führungsebene von Universitäten und Fachhochschulen (Präsidenten, Kanzler, Dekane sowie Professoren) geführt. Im Vordergrund steht dabei die Frage, wie die Hochschulen intern auf die (politisch vorgegebenen) Veränderungen von außen reagieren, sowohl auf Ebene der Hochschulleitung als auch auf Ebene einzelner Fakultäten. Insbesondere wird der Frage nachgegangen, ob sich gleichzeitig die Identität von Hochschulen als Organisationen verändert und ob Universitäten und Fachhochschulen im Zuge der Reformen neue Hochschultypen ausbilden. Die wesentlichen Aussagen der empirischen Untersuchung fließen als Zitate in den Fließtext ein und eröffnen so einen interessanten Blick hinter die Kulissen deutscher Hochschulen. Die Studie wurde 2011 mit dem Ulrich-Teichler-Preis für die beste Dissertation in der Hochschulforschung ausgezeichnet.

Heinze, Thomas; Krücken, Georg (Hrsg.): Institutionelle Erneuerungsfähigkeit der Forschung. Wiesbaden: Springer, 2012, ISBN 978-3-531-18469-2, 262 Seiten

Die Fähigkeit von Hochschulen und außeruniversitären Forschungseinrichtungen, neue Forschungsfelder zu erschließen sowie zur Erzeugung neuen Wissens und zur Ausbildung des wissenschaftlichen Nachwuchses sind zwei Aspekte des in diesem Aufsatzband behandelten Themas der Erneuerungsfähigkeit von Forschung. Das Buch verfolgt dieses Thema nicht aus der Perspektive der Disziplinen, sondern aus der Sicht der Organisationsforschung. Mit Forschung befasste Organisationen sollen neue Forschungsfelder etablieren und voranbringen, neues Wissen generieren und wissenschaftlichen Nachwuchs heranziehen, sowie schließlich auch zu Trägern des institutionellen Wandels der Forschung im Zuge gesellschaftlicher Veränderungen werden. Der Sammelband mit acht vorwiegend sozialwissenschaftlich ausgerichteten Beiträgen erschließt, auch mithilfe von Rückblicken, das Thema Erneuerungsfähigkeit für die Organisationsforschung. Das Buch wendet sich an Wissenschafts- und Hochschulforscher, Hochschulmanager und Soziologen mit den Schwerpunkten Hochschul- und Organisationsforschung.

Hinweise für Autoren

Konzept:

Die Zeitschrift „Beiträge zur Hochschulforschung“ bietet Hochschulforschern und Akteuren im Hochschulbereich die Möglichkeit zur Erstveröffentlichung von Artikeln, die wichtige Entwicklungen im Hochschulbereich aus unterschiedlichen methodischen und disziplinären Perspektiven behandeln. Dabei wird ein Gleichgewicht zwischen quantitativen und qualitativen empirischen Analysen, Vergleichsstudien und Überblicksartikeln angestrebt.

Eingereichte Artikel sollten klar und verständlich formuliert, übersichtlich gegliedert sowie an ein Lesepublikum aus unterschiedlichen Disziplinen mit wissenschaftlichem und praxisbezogenem Erwartungshorizont gerichtet sein.

Review-Verfahren:

Wie für eine wissenschaftliche Zeitschrift üblich, durchlaufen alle eingereichten Manuskripte eine zweifache Begutachtung durch anonyme Sachverständige (double blind) innerhalb und außerhalb des Instituts. Dabei kommen je nach Ausrichtung des Artikels folgende Kriterien zum Tragen: Relevanz des Themas, Berücksichtigung des hochschulpolitischen Kontexts, Praxisbezug, theoretische und methodische Fundierung, Qualität der Daten und empirischen Analysen, Berücksichtigung der relevanten Literatur, klare Argumentation und Verständlichkeit für ein interdisziplinäres Publikum. Die Autoren werden über das Ergebnis schriftlich informiert und erhalten gegebenenfalls Hinweise zur Überarbeitung. Die redaktionelle Betreuung der Zeitschrift liegt bei Mitarbeitern des Instituts.

Umfang und Form der eingereichten Manuskripte:

Manuskripte sollten bevorzugt per E-Mail eingereicht werden und einen Umfang von 20 Seiten/50.000 Zeichen mit Leerzeichen nicht überschreiten (Zeilenabstand 1,5, Arial 11). Ergänzend sollten je ein Abstract (maximal 1000 Zeichen mit Leerzeichen) in deutscher und in englischer Sprache sowie Anschrift und Angaben zur beruflichen Funktion des Autors beigefügt sein. Die Druckfassung wird extern von einem Grafiker erstellt. Weitere wichtige Hinweise zur Gestaltung der Manuskripte finden Sie auf unserer Homepage www.ihf.bayern.de unter Publikationen.

Kontakt:

Dr. Lydia Hartwig

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF)

Prinzregentenstraße 24

80538 München

E-Mail: Hartwig@ihf.bayern.de

Aus dem Inhalt

Ester Höhle, Anna Katharina Jacob, Ulrich Teichler: Das Paradies nebenan?
Zur Situation von Wissenschaftlerinnen und Wissenschaftlern an außer-
universitären Forschungseinrichtungen und Universitäten in Deutschland

Gero Lenhardt, Robert D. Reisz, Manfred Stock: Überlebenschancen
privater und öffentlicher Hochschulen im Ländervergleich

Werner Marx, Lutz Bornmann: Der Journal Impact Factor:
Aussagekraft, Grenzen und Alternativen in der Forschungsevaluation

Harry Müller: Zitationen als Grundlage von Forschungsleistungsrankings –
Konzeptionelle Überlegungen am Beispiel der Betriebswirtschaftslehre

Beiträge zur HOCHSCHULFORSCHUNG

3 | 2012

Thema: Qualitätssicherung in Lehre und Forschung

Mittag/Mutz/Daniel: Anforderungen an Qualitätssicherungsinstrumente

Löscher: Qualitätssicherung in Großbritannien

Gülker: Wissenschaft aus Kalkül

Banscherus/Reiß: Das sächsische Hochschulrating

Beiträge zur HOCHSCHULFORSCHUNG

3 | 2012

Thema: Qualitätssicherung in Lehre und Forschung

Mittag/Mutz/Daniel: Anforderungen an Qualitätssicherungsinstrumente

Löscher: Qualitätssicherung in Großbritannien

Gülker: Wissenschaft aus Kalkül

Banscherus/Reiß: Das sächsische Hochschulrating

Impressum

Beiträge zur Hochschulforschung

erscheinen viermal im Jahr

ISSN 0171-645X

Herausgeber: Bayerisches Staatsinstitut für Hochschulforschung
und Hochschulplanung, Prinzregentenstraße 24, 80538 München
Tel.: 089/2 1234-405, Fax: 089/2 1234-450

E-Mail: Sekretariat@ihf.bayern.de

Internet: <http://www.ihf.bayern.de>

Herausgeberbeirat:

Mdgt. a. D. Jürgen Großkreutz, Bayerisches Staatsministerium für Wissenschaft,
Forschung und Kunst, München

Dr. Lydia Hartwig, Bayerisches Staatsinstitut für Hochschulforschung und
Hochschulplanung, München

Professor Dr. Dorothea Jansen, Deutsche Hochschule für Verwaltungswissen-
schaften, Speyer

Professor Dr. Dr. h. c. Hans-Ulrich Küpper, Bayerisches Staatsinstitut für Hochschul-
forschung und Hochschulplanung und Ludwig-Maximilians-Universität, München

Thomas May, Wissenschaftsrat, Köln

Professor Rosalind Pritchard, AcSS, University of Ulster, United Kingdom

Redaktion: Dr. Lydia Hartwig (V.i.S.d.P.)

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung

E-Mail: Hartwig@ihf.bayern.de

Die abgedruckten Beiträge geben die Meinung der Verfasser wieder.

Graphische Gestaltung: Haak & Nakat, München

Satz: Dr. Ulrich Scharmer, München

Druck: Steinmeier, Deiningen

Ausrichtung, Themenspektrum und Zielgruppen

Die „Beiträge zur Hochschulforschung“ sind eine der führenden wissenschaftlichen Zeitschriften im Bereich der Hochschulforschung im deutschen Sprachraum. Sie zeichnen sich durch hohe Qualitätsstandards, ein breites Themenspektrum und eine große Reichweite aus. Kennzeichnend sind zudem die Verbindung von Wissenschaftlichkeit und Relevanz für die Praxis sowie die Vielfalt der Disziplinen und Zugänge. Dabei können die „Beiträge“ auf eine lange Tradition zurückblicken. Die Zeitschrift erscheint seit ihrer Gründung 1979 viermal im Jahr und publiziert Artikel zu Veränderungen in Universitäten, Fachhochschulen und anderen Einrichtungen des tertiären Bildungsbereichs sowie Entwicklungen in Hochschul- und Wissenschaftspolitik in nationaler und internationaler Perspektive.

Wichtige Themenbereiche sind:

- Strukturen der Hochschulen,
- Steuerung und Optimierung von Hochschulprozessen,
- Hochschulfinanzierung,
- Qualitätssicherung und Leistungsmessung,
- Studium und Studierende, Umsetzung des Bologna-Prozesses,
- Übergänge zwischen Schule, Hochschule und Arbeitsmarkt,
- Forschung und wissenschaftlicher Nachwuchs, akademische Karrieren,
- Frauen in Hochschulen und Wissenschaft,
- Wissenschaft und Wirtschaft,
- International vergleichende Hochschulforschung.

Die Zeitschrift veröffentlicht quantitative und qualitative empirische Analysen, Vergleichsstudien und Überblicksartikel, die ein anonymes Peer Review-Verfahren durchlaufen haben. Sie bietet die Möglichkeit zum Austausch von Forschungsergebnissen und stellt ein Forum für Hochschulforscher und Experten aus der Praxis dar. Zwei Ausgaben pro Jahr sind in der Regel einem aktuellen hochschulpolitischen Thema gewidmet, die beiden anderen sind inhaltlich nicht festgelegt. Es besteht die Möglichkeit, Aufsätze in deutscher und englischer Sprache einzureichen. Hinweise für Autoren befinden sich auf der letzten Seite.

Die „Beiträge“ richten sich an Wissenschaftler, die sich mit Fragen des Hochschulwesens und seiner Entwicklung befassen, aber auch an politische Entscheidungsträger, Hochschulleitungen, Mitarbeiter in Hochschulverwaltungen, Ministerien sowie Wissenschafts- und Hochschulorganisationen.

Inhalt

Editorial	4
Abstracts	6
Sandra Mittag, Rüdiger Mutz, Hans-Dieter Daniel: Anforderungen an Qualitätssicherungsinstrumente für Lehre und Studium an Hochschulen – Ergebnisse einer Meta-Evaluation an der ETH Zürich	8
Anett Löscher: Qualitätssicherung in Großbritannien durch die Quality Assurance Agency – Ein Überblick	32
Silke Gülker: Wissenschaft aus Kalkül? Wissenschaftsinterne Antizipation von Wettbewerbsmechanismen im Rahmen von Evaluationsverfahren	50
Ulf Banscheraus, Mirjam Reiß: Das sächsische Hochschulrating als Beispiel für einen evidenzbasierten Ansatz in der Hochschulentwicklungsplanung – Ein Werkstattbericht	72
Buchvorstellungen	92
Hinweise für Autoren	94

Editorial

Welche Instrumente der Qualitätssicherung werden in Lehre und Forschung eingesetzt, in welchen hochschulpolitischen Zusammenhängen stehen sie, welche Zielrichtungen verfolgen sie, und welche Wirkungen bringen sie hervor? Diesen Fragen gehen die Autoren in dem vor Ihnen liegenden Themenheft der „Beiträge zur Hochschulforschung“ nach. Dabei werden jeweils unterschiedliche Ebenen in den Blick genommen: die einer Hochschule, einer Wissenschaftsorganisation, eines Bundeslandes und eines Staates.

Sandra Mittag, Rüdiger Mutz und Hans-Dieter Daniel stellen zentrale Ergebnisse einer Meta-Evaluation der Qualitätssicherungsinstrumente der ETH Zürich im Bereich Lehre und Studium vor. Das Qualitätssicherungssystem der ETH Zürich beruht auf den drei Eckpfeilern Lehrveranstaltungsbeurteilung, Absolventenbefragung und Departements-evaluation. Die Autoren zeigen, dass diese Instrumente erfolgreich eingesetzt werden und breit akzeptiert sind, dass es jedoch auch Optimierungspotentiale gibt, so z. B. bei der Verzahnung der einzelnen Bausteine und der Entwicklung eines in sich kohärenten Qualitätssicherungssystems.

In Großbritannien obliegt die Aufgabe der Qualitätssicherung in der Lehre der Quality Assurance Agency (QAA). Anett Löscher gibt aus der Innensicht einer Mitarbeiterin der QAA einen Überblick über die verschiedenen Begutachtungsmethoden, die sowohl den hochschulpolitischen Rahmenbedingungen der vier Länder des Vereinigten Königreichs (England, Wales, Schottland, Nordirland) als auch den Besonderheiten der unterschiedlichen Bildungsanbieter Rechnung tragen. Sie stellt insbesondere den UK Quality Code for Higher Education vor, der Grundlage für die Begutachtungen der QAA ist, und zeigt dabei, wie sich ein Wechsel der Perspektive von der Beurteilung der Lehre hin zum Urteil durch die Studierenden vollzieht.

Der Artikel von Silke Gülker widmet sich der Frage, ob und wie Evaluationsverfahren gängige Kooperationsformen in der Wissenschaft verändern. Durch die Analyse zweier Evaluationsverfahren in der Forschung, der Institute der Leibniz-Gemeinschaft in Deutschland und der Research Assessment Exercises in Großbritannien, zeigt die Autorin, dass die jeweiligen Evaluationsverfahren zwar ein strategisches Verhalten der beteiligten Wissenschaftler begünstigen, die Normen der Wissensgesellschaft aber weiterhin ihre Bedeutung behalten.

Eine breite Fächerevaluation auf Landesebene für die Zwecke der Hochschulplanung ist Gegenstand des Artikels von Ulf Banscheraus und Mirjam Reiß, die das sächsische Hochschulrating vorstellen. Sie beschreiben das mehrstufige Bewertungsverfahren ausgewählter Fächer an den Universitäten und Fachhochschulen des Freistaates

Sachsen, welches eine umfangreiche Datenauswertung mit einer qualitativen Begutachtung durch externe Expertinnen und Experten verband und die Landeshochschulplanung auf eine evidenzbasierte Grundlage stellen sollte. Auf der Basis der Ergebnisse wurden Vorschläge zur stärkeren Abstimmung des Fächerangebots zwischen den Hochschulen sowie zur Konzentration des Studienangebots auf bestimmte Standorte formuliert.

Lydia Hartwig, Sandra Mittag

Sandra Mittag, Rüdiger Mutz, Hans-Dieter Daniel: Requirements for quality assurance instruments of teaching and learning – Results of a meta-evaluation at ETH Zurich

The article presents results of a meta-evaluation of quality assurance instruments of teaching and learning at ETH Zurich. Cornerstones of quality assurance are student evaluations of teaching, graduate surveys, and department evaluations (including self-evaluation and peer review). The study indicates that ETH Zurich has established and widely accepted instruments for quality assurance in teaching and learning. Nevertheless, all of the evaluated instruments have optimisation potential. The survey instruments, for example, have not yet made the shift from teaching to learning, and the different instruments do not yet form an integral part of a comprehensive quality management system. The results of this article will also be published in a book-length manuscript (Mittag/Mutz/Daniel 2012).

Quality assurance in Great Britain by the Quality Assurance Agency – An overview

The Quality Assurance Agency for Higher Education (QAA) checks whether and how British higher education providers (public and private universities, university colleges and colleges, and further education colleges) set, maintain and safeguard academic standards, and enhance the quality of the learning opportunities they offer. QAA has developed review methods for all higher education provision that leads to a British higher education award. These methods take into consideration both the respective contexts for higher education in the UK's four nations, and the type of provider that is checked through the review method in question. This article offers an overview of QAA's review methods. It also introduces the UK Quality Code for Higher Education, which is the nationally agreed reference point for quality assurance in British higher education.

Silke Gülker: Scholarship driven by calculation? Science internal anticipation of competition mechanisms within evaluation procedures

To what extent are evaluation procedures capable of shaking up established mechanisms of coordination within science and research? The article addresses this question on a conceptual and an empirical level. While science has traditionally accepted competition (for the best ideas) and community (cooperation based on collegiality) as a couple, hierarchical decision-making plays an inferior role. Analysing evaluation procedures of member-institutes of the Leibniz-Association (Leibniz-Gemeinschaft) in Germany and of the Research Assessment Exercises in the United Kingdom, we show

that the procedures interfere indeed with these accepted mechanisms, but they do not fundamentally challenge them. Evaluation procedures provoke strategic behaviour of scientists, but at the same time, norms of the scientific community maintain their decisive importance. The preparation of evaluation procedures requires new decision-making processes, which are resolved hierarchically.

Ulf Banscherus, Mirjam Reiß: The Saxon Higher Education Rating as an example of an evidence-based approach in higher education development planning – A workshop report

Against the background of demographic change, the Free State of Saxony implemented the Saxon Higher Education Rating, a pilot project which aimed to create the basis for evidence-based governmental higher education development planning. For this purpose, a multi-stage evaluation process of selected fields of study at higher education institutions was developed which combined an extensive quantitative data analysis with a qualitative review by external experts. The results both contain overarching strategic recommendations for the policy of higher education as well as proposals for greater coordination among higher education institutions concerning the range of courses offered including the concentration of study courses to specific locations. Due to the high transparency of the process and the extensive opportunities for participation, the results were largely accepted by the higher education institutes and formed the basis for the continuation of the Saxon Higher Education Development Plan.

Anforderungen an Qualitätssicherungsinstrumente für Lehre und Studium an Hochschulen – Ergebnisse einer Meta-Evaluation an der ETH Zürich

Sandra Mittag, Rüdiger Mutz, Hans-Dieter Daniel

Der vorliegende Beitrag stellt die zentralen Ergebnisse einer Meta-Evaluation der Qualitätssicherungsinstrumente der ETH Zürich im Bereich Lehre und Studium vor¹. Die Eckpfeiler der Qualitätssicherung sind die Instrumente Lehrveranstaltungsbeurteilung, Absolventenbefragung sowie die Departementsevaluation mit Selbstevaluation und Peer Review. Es hat sich gezeigt, dass die ETH Zürich über etablierte Instrumente der Qualitätssicherung im Bereich Lehre und Studium verfügt, die in der Hochschule weitestgehend akzeptiert sind. Alle Instrumente verfügen jedoch über Optimierungspotentiale. So wurde beispielsweise noch nicht der Perspektivenwechsel vom Lehren zum Lernen vollzogen. Darüber hinaus sind die verschiedenen Instrumente noch nicht integraler Bestandteil eines umfassenden Qualitätsmanagementsystems.

1 Einleitung

Die Schweiz nimmt eine Vorreiterrolle für die institutionelle Qualitätssicherung an Hochschulen ein. Neben den Qualitätssicherungssystemen an den Hochschulen selbst kommt hierbei dem Organ für Akkreditierung und Qualitätssicherung der Schweizerischen Hochschulen (OAQ) eine besondere Bedeutung zu.

In der Schweiz überprüft das OAQ im Auftrag des Staatssekretariats für Bildung und Forschung (SBF) periodisch im Rahmen institutioneller Verfahren, ob die Qualitätssicherungssysteme der Universitäten Mindestanforderungen erfüllen, die der internationalen Praxis entsprechen und sich auf die „Standards and Guidelines for internal quality assurance within higher education institutions“ der European Association for Quality Assurance in Higher Education (ENQA) in der Fassung vom Februar 2005 stützen.²

¹ Die Ergebnisse dieses Artikels beruhen auf einer Studie, die in ausführlicher Fassung 2012 erscheint (Mittag/Mutz/Daniel 2012).

² Vgl. die Qualitätssicherungs-Richtlinien der Schweizerische Universitätskonferenz, <http://www.oaq.ch/pub/de/documents/D-443-06A-Quali-RL-VO.pdf> (Zugriff: 01.06.2012)

Die Überprüfung der Qualitätssicherungssysteme erfolgt in Form von Quality Audits³. Die Qualitätsprüfung ist gemäß Universitätsförderungsgesetz (UFG) vom 8.10.1999 obligatorisch für alle kantonalen Universitäten und Voraussetzung für eine Finanzhilfe des Bundes. An den Quality Audits, die in den Jahren 2003/2004 und 2007/2008 durchgeführt wurden, hat sich die ETH Zürich als eine vom Bund finanzierte Universität freiwillig beteiligt.

Mit der Durchführung des ersten Quality Audit 2003/2004 des OAO haben die Hochschulen begonnen, umfassende Qualitätssicherungssysteme aufzubauen. In Vorbereitung auf das Quality Audit 2007/2008 wurde das Qualitätssicherungssystem der ETH Zürich für den Bereich Lehre einer umfassenden Meta-Evaluation unterzogen. Es stützt sich auf die vier Instrumente Lehrveranstaltungsbeurteilung, Absolventenbefragung (so genannte Diplomiertenbefragung), Selbstevaluation und Peer Review⁴. Die folgenden Fragen standen im Vordergrund der Untersuchung:

- Wie steht es um die Qualität der Instrumente, die an der ETH Zürich für die Qualitätssicherung und -entwicklung von Lehre und Studium eingesetzt werden? Entsprechen sie den wissenschaftlichen Standards?
- Welche Rolle spielt der Bereich Lehre und Studium in den Selbstevaluationen und Peer Reviews im Rahmen der Departementsevaluationen? Welche Themen werden fokussiert?

Darüber hinaus interessierte die Frage, wie die Absolventen der ETH Zürich rückblickend ihr Studium bewerten. Die Studie gründet auf den Ergebnissen des Projekts „Metaevaluation der Qualitätssicherungsinstrumente der ETH Zürich im Bereich Lehre und Studium“, das im Auftrag des Rektors der ETH Zürich im Jahr 2007 durchgeführt wurde und das eine wichtige Rolle im Kontext des Quality-Audits des OAO in den Jahren 2007/2008 gespielt hat (Organ für Akkreditierung und Qualitätssicherung 2008). Im Folgenden werden zunächst die vier Instrumente in ihren Zielsetzungen und Abläufen vorgestellt. Anschließend wird auf die Methoden und die zentralen Ergebnisse eingegangen. Der Beitrag schließt mit einem Fazit.

³ Mit einem Quality Audit wird das Qualitätssicherungssystem einer Hochschule überprüft. Es besteht aus einem zweistufigen Verfahren: erstens der Selbstbeurteilung der Universität oder universitären Institution und zweitens der Visite durch unabhängige Qualitätssicherungsexperten. Das OAO verfasst Schlussberichte, die im Internet publiziert werden.

⁴ Selbstevaluation und Peer Review sind die beiden Phasen der Departementsevaluation. Sie wird an der ETH Zürich auch für die Qualitätssicherung und -entwicklung von Forschung und Dienstleistung eingesetzt. Als „Peer Review“ wird die Bewertung durch eine internationale Expertenkommission bezeichnet.

2 Instrumente der Qualitätssicherung an der ETH Zürich

Die ETH Zürich ist eine der beiden Eidgenössisch Technischen Hochschulen der Schweiz. Im Wintersemester 2011/2012 waren an der ETH Zürich über 17.000 Studierende eingeschrieben; mehr als 400 Professorinnen und Professoren lehren und forschen an der ETH Zürich. Die zentralen Qualitätssicherungsinstrumente der ETH Zürich im Bereich Lehre und Studium sind die Lehrveranstaltungsbeurteilung, die Absolventenbefragung (so genannte Diplomiertenbefragung) sowie die Departementsevaluation, die aus Selbstevaluation und Peer Review besteht. Da der Selbstevaluation an der ETH Zürich eine große Bedeutung beigemessen wird, wird sie als ein eigenständiges Instrument angesehen.

2.1 Lehrveranstaltungsbeurteilung

Die Lehrveranstaltungsbeurteilung durch Studierende, die an der ETH Zürich als Unterrichtsbeurteilung (UB) bezeichnet wird, ist ein zentrales Element des Qualitätssicherungssystems von Lehre und Studium an der ETH (Niedermann/Hugentobler/Osterwalder, 2003; Alean-Kirkpatrick/Hänni/Lutz 1997). Ziel der Lehrveranstaltungsbeurteilung ist es, die Lehre hinsichtlich didaktischer, inhaltlicher, curricularer und organisatorischer Aspekte kontinuierlich zu verbessern. Dem jeweiligen Departement obliegt die Pflicht, die Qualität seines gesamten Lehrangebots zu sichern und zu verbessern. Die Lehrveranstaltungsbeurteilung soll den für Lehre verantwortlichen Personen in den Departementen sowie dem Rektor Einblick in die Lehrleistung der Dozierenden aus Sicht der Studierenden geben. Die systematische Lehrveranstaltungsbeurteilung findet seit Anfang der 1990er Jahre statt; in der Regel wird sie alle drei Semester durchgeführt. An einem Departement werden jeweils alle Dozierenden und alle Lehrveranstaltungen in die Befragung durch die Studierenden einbezogen.

Die Fragebögen zur Lehrveranstaltungsbeurteilung bestehen aus sechs Teilen:

1. Angaben zur Lehrveranstaltung,
2. personenbezogene Fragen an die Studierenden (Geschlecht u. a.),
3. Fragen des Rektors,
4. Fragen des Studiengangs,
5. Fragen der Dozentin/des Dozenten und
6. zusätzlicher Kommentar.

2.2 Absolventenbefragung („Diplomiertenbefragung“)

Die Absolventenbefragung wird seit 1997 an der ETH Zürich im Rahmen der Departementsevaluation von der Direktion Finanzen und Controlling der ETH Zürich, Abteilung Controlling, durchgeführt. Die Absolventen des zu evaluierenden Departements wer-

den a) zu ihrer gegenwärtigen Beschäftigungssituation und b) rückblickend zu Lehre und Studium an der ETH Zürich befragt. Die Absolventenbefragung hat die Zielsetzung, „Aufschluss darüber zu verschaffen, wie die ehemaligen Studierenden, die nun im Berufsleben oder in der Forschung stehen, ihre an der ETH erfahrene Ausbildung beurteilen. Die Rückmeldungen fließen auf verschiedenen Ebenen in den Erneuerungsprozess der Studienprogramme ein“ (Niedermann/Hugentobler/Osterwalder 2003, S. 18). In der Regel werden die Absolventen zwei, drei und vier Jahre nach Abschluss ihres Studiums an der ETH Zürich schriftlich befragt. Dabei wird ein Standardfragebogen verwendet, um eine Vergleichbarkeit zu gewährleisten; in Absprache mit dem jeweiligen Departement kann der Fragebogen jedoch angepasst werden.⁵ Die folgenden drei Themenbereiche werden behandelt:

1. Allgemeines: Fragen zur Person, zum Studium, zu weiteren Qualifikationen nach dem Studium (z. B. Promotion) und zum Bedarf an Weiterbildung.
2. Berufstätigkeit: Erwerbstätigkeit, Adäquatheit der Stelle, Unternehmen, in dem die befragte Person tätig ist, Position, Verantwortungsbereich sowie Einkommen.
3. Studium im Rückblick: Fragen zu Aktualität der Studieninhalte, Umfang von Fach- und Methodenwissen sowie Rahmenkompetenzen, Beitrag der ETH Zürich an der persönlichen Entwicklung verschiedener Kompetenzbereiche (Mutz/Daniel 2008). Darüber hinaus wird um Bewertungen der Aufteilung der verschiedenen Lehrveranstaltungstypen gebeten (z. B. zu viele oder zu wenige Vorlesungen im Grund- oder Fachstudium). Der letzte Teil des Bogens beinhaltet offene Fragen zu Lehrveranstaltungen, die inhaltlich oder methodisch als besonders wichtig oder unwichtig in Erinnerung geblieben sind bzw. gefehlt haben. Darüber hinaus sollen auch akademische Lehrer genannt werden, die besonders in Erinnerung geblieben sind. Am Schluss wird gefragt, ob die befragte Person gegebenenfalls wieder ein Studium ergreifen würde, ob sie dasselbe Fach studieren würde und dies erneut an der ETH Zürich.

Darüber hinaus werden die Absolventen der ETH Zürich unmittelbar nach ihrem Studienabschluss im Rahmen der so genannten Beschäftigungsstatistik über ihre weiteren beruflichen Vorstellungen und Möglichkeiten befragt. Die Fragebögen werden zeitnah nach Studienabschluss durch das Rektorat verschickt und von der Abteilung Finanzen und Controlling der ETH Zürich einmal pro Jahr ausgewertet. Die Ergebnisse dieser Befragungen sollen der Hochschulleitung, den Departementen und den Studierenden Hinweise über die Berufschancen und den Verbleib der Absolventen geben.

⁵ Musterfragebögen finden sich unter <http://www.let.ethz.ch/unterricht/evaluation/dienstleistungen/fragebogen/fragebogenliste> (Zugriff: 01.06.2012).

2.3 Departementsevaluation

An der ETH Zürich werden seit 1989 Departementsevaluationen als mehrstufige Evaluationsverfahren durchgeführt. Die erste Phase bildet die Selbstevaluation, auf die das Peer Review folgt (vgl. Kübler 2003; Ulich 1999)⁶. Zunächst war der ETH-Rat Auftraggeber der Evaluationen, seit Anfang der 1990er Jahre werden sie im Auftrag der Hochschulleitung und zum Teil in Absprache mit der ETH Lausanne durchgeführt. Während anfänglich der Fokus der Evaluationsverfahren auf der Forschung lag, wurden bald auch Lehre und Dienstleistungen einbezogen. Im Handbuch zum Qualitätsmanagement wird die Selbstevaluation aufgrund ihrer großen Bedeutung, die sie im Rahmen des Qualitätsmanagementsystems an der ETH Zürich einnimmt, als gesondertes Instrument aufgeführt.

2.3.1 Selbstevaluation

Die Selbstevaluation ist die Vorbereitung eines Departements auf die Bewertung durch ein internationales Expertenteam (Peer Review). Es wird eine umfangreiche Dokumentation erstellt, in deren Rahmen sich das Departement der Expertenkommission präsentiert und Auskunft über festgelegte Inhalte gibt (insbesondere über die Organisation, den Studienplan, Stärken und Schwächen, Mission Statement hinsichtlich Lehre, Forschung und Dienstleistung). Es gibt keinen detaillierten Frageleitfaden für die Dokumentation der Departemente; spezifische Fragestellungen werden bei jeder Evaluation, angepasst an das zu evaluierende Departement, neu formuliert. Die Departemente sind in der Ausgestaltung ihrer Dokumentationen und damit auch bei der Gewichtung der verschiedenen Inhalte relativ frei.

2.3.2 Peer Review

Auf die Phase der Selbstevaluation folgt das Peer Review, in dessen Rahmen der Besuch des internationalen Expertenteams stattfindet und ein Evaluationsbericht für die Hochschulleitung verfasst wird. Bis 1999 wurden sämtliche Departemente von einer internationalen Expertenkommission evaluiert; 1998 wurde mit der zweiten, 2006 mit der dritten Runde begonnen. Die Departemente werden in einem Turnus von sechs bis neun Jahren evaluiert.

Der Evaluationsbericht der Experten umfasst die folgenden Inhalte:

- Zusammenfassung der Empfehlungen,
- Übersicht des Departements,
- generelle Eindrücke, Verbesserungspotential,

⁶ Für eine Übersicht zu den mehrstufigen Evaluationsverfahren siehe Mittag/Bornmann/Daniel (2003) und Mittag (2006).

- Lehre,
- Forschung,
- Infrastruktur und Dienstleistungen,
- Beurteilung und Empfehlungen zu einzelnen Instituten und Professuren,
- Aussichten.

Die Experten erhalten von Seiten der Hochschulleitung auf das jeweils zu evaluierende Departement bezogene spezifische Fragestellungen. Sie haben jedoch in ihren Evaluationsberichten ebenso wie die Departemente bei der Erstellung der Dokumentationen bezüglich der inhaltlichen Ausrichtung und der Gewichtung der verschiedenen Themen große Freiheiten.

3 Methoden der Meta-Evaluation

3.1 Sekundärstatistische Analysen

Für die Studie wurden die Daten aus a) allen Lehrveranstaltungsbeurteilungen, die zwischen dem Wintersemester 2003/2004 und dem Sommersemester 2006 durchgeführt wurden, und b) allen Absolventenbefragungen, die zwischen 1997 und 2006 durchgeführt wurden, sekundärstatistisch ausgewertet. Darüber hinaus sind Daten aus Lehrveranstaltungsbeurteilungen von Departementen einbezogen worden, die zwischen 1997 und 2007 eine Selbstdokumentation erstellt haben und extern evaluiert wurden (Peer Review). Mit der Unterstützung der Abteilung Controlling der Direktion Finanzen und Controlling wurden für jede Lehrveranstaltung die Teilnehmerzahlen ermittelt, um Rücklaufquoten zu generieren. Diese Datenbasis ergibt sich aus dem Vorläuferprojekt „Roadmap: Maßnahmen zur Verbesserung von Lehre und Studium an der ETH Zürich“ (2006-2008; Abd-el-Razik/Eugster/Grund/Lutz/Mittag et al. 2007).

Im Roadmap-Projekt wurde zudem im Januar/Februar 2007 eine Online-Befragung aller Dozierenden an der ETH Zürich durchgeführt. Die Personengruppe der Dozierenden umfasst alle ordentlichen und außerordentlichen Professoren, Assistenzprofessoren, Privatdozierende, Dozierende (ehemals Lehrbeauftragte), Titularprofessoren der ETH Zürich sowie weitere Titularprofessoren. Insgesamt wurden 2 171 Dozierende per E-Mail angeschrieben, von denen 2 075 die Adressatin bzw. den Adressaten erreichten. 473 Dozierende beteiligten sich an der Befragung; dies entspricht einer Rücklaufquote von 23 Prozent. Je nach Fragestellung wurden in die Meta-Evaluation die Aussagen aller Dozierenden (n=473) bzw. nur die Aussagen der ordentlichen und außerordentlichen Professoren sowie der Assistenz- und Titularprofessoren (n=145) zu diesen Fragen einbezogen.

3.2 Literaturanalyse

Um die Verfahren der Lehrveranstaltungsbeurteilung und der Absolventenbefragung an der ETH Zürich vor dem Hintergrund des aktuellen wissenschaftlichen Forschungsstandes bewerten zu können, wurde eine Analyse der wissenschaftlichen Literatur zu Lehrveranstaltungsbeurteilungen und Absolventenbefragungen durchgeführt. Da die Anzahl von Einzelartikeln insbesondere zur Lehrveranstaltungsbeurteilung ein Ausmaß erreicht, das im Rahmen eines qualitativen Literaturreviews nicht mehr zu bewältigen ist, orientiert sich die Literaturanalyse an zentralen Übersichtsarbeiten, Review-Artikeln und Monographien, die in den letzten zehn Jahren veröffentlicht wurden. Die recherchierte Literatur wurde im Hinblick auf international gültige fachliche Standards von Lehrveranstaltungsbeurteilungen und Absolventenbefragungen ausgewertet. In den letzten 20 Jahren wurde eine Reihe von Übersichtsarbeiten publiziert, die den jeweils aktuellen internationalen Forschungsstand im Bereich der studentischen Lehrveranstaltungsbeurteilung (Lehrevaluation, course evaluation, student evaluation of teaching or instruction, student rating of instruction) zusammenfassen und Hinweise geben, wie Lehrveranstaltungsbeurteilungen zu gestalten sind. Für die Zusammenfassung des Forschungsstands zu verschiedenen Themenbereichen der Lehrevaluation siehe Mittag/Mutz/Daniel 2012.

3.3 Dokumentenanalyse

Im Rahmen der Dokumentenanalyse wurden 67 Dokumente aus insgesamt 20 Departementsevaluationen ausgewertet, die zwischen 1997 und 2006 durchgeführt wurden. Für kein Verfahren lagen dem Projektteam alle fünf Dokumententypen – Dokumentation (Selbstevaluation), Evaluationsbericht der Experten, Stellungnahme des Departements, Stellungnahme der Hochschulleitung und Umsetzungsbericht – vor. Zugleich kann es für ein Evaluationsverfahren mehrere Dokumente desselben Dokumententyps geben, da sich die Berichte teilweise aus mehreren Teilberichten zusammensetzen. Die Dokumente verteilen sich über die verschiedenen Evaluationsverfahren und Dokumententypen gemäß Tabelle 1.

Tabelle 1: Anzahl der Verfahren und Anzahl der analysierten Dokumente nach Dokumenttyp (1997–2006)

Dokumenttyp	Anzahl Evaluationsverfahren	Anzahl analysierter Dokumente
Dokumentation (Selbstevaluation)	10	16
Evaluationsbericht (Peer Review)	18	20
Stellungnahme Departement	13	17
Stellungnahme Hochschulleitung	13	14
Umsetzungsbericht	3	0
Insgesamt	20	67 ^a

^aIn jedem Evaluationsverfahren treten mehrere Dokumenttypen auf.

Aufgrund der geringen Anzahl Dokumente des Typs „Umsetzungsbericht“ wurden diese nicht in die Analyse einbezogen. Alle anderen Dokumente wurden themenanalytisch ausgewertet. Hierfür ist ein Kategoriensystem entwickelt worden, das die genannten Fragestellungen berücksichtigt und die verschiedenen Auswertungsebenen einbezieht (Mayring 2008). Das Kategoriensystem wurde zu Beginn der Auswertung einem Pretest unterzogen und entsprechend adaptiert. Bei der Analyse der Dokumentationen der Departemente (Selbstevaluation) wurden drei Ebenen unterschieden: die Ebene Departement/Abteilung, die Ebene Institute/Labore/Forschungsgruppen sowie die Ebene Professoren⁷. Die Anzahl der Analyseeinheiten auf der Ebene der Departemente entspricht nicht der Anzahl der involvierten Departemente nach der heutigen Struktur, da in eine Evaluation mehrere Abteilungen einbezogen sein können (welche z. B. später zu einem Departement zusammengefasst wurden). Somit beruht die Analyse auf insgesamt 16 Bereichen, Abteilungen bzw. Departementen in offiziell zehn Verfahren, aus denen die Dokumentationen (Selbstevaluation) vorlagen (siehe Tabelle 2). Insgesamt 59 Institute/Labore/Forschungsgruppen und 242 Professuren wurden in diesen 16 Berichten dokumentiert. Innerhalb einer Ebene konnten mehrere Textstellen identifiziert werden, die für die Untersuchungsfragen relevant sind. Daher gibt es auf jeder Ebene ein Vielfaches an kodierten Textstellen (siehe dritte und vierte Spalte von Tabelle 2). Insgesamt wurden in den Dokumentationen 1 436 Textstellen kodiert.

Die Dokumente wurden dem Projektteam in zwei Phasen zur Verfügung gestellt (Ende März/Anfang April 2007 und Anfang August 2007). Die Dokumente aus der ersten Phase (1. Erhebungswelle) wurden bei einigen Fragestellungen detaillierter untersucht als die Dokumente der zweiten Phase (siehe jeweils die dritte Spalte der beiden folgenden Tabellen). Bei der Ergebnisdarstellung finden sich entsprechende Verweise.

⁷ In Anlehnung an die Darstellung in den Dokumenten wurden teilweise mehrere Professoren zu einer Professur zusammengefasst.

Tabelle 2: Anzahl Analyseeinheiten bei den Dokumentationen der Departemente (Selbstevaluation)

Analyseeinheit	Anzahl Einheiten je Ebene	Anzahl kodierter Textstellen je Ebene (1. Erhebungswelle; n = 1388)	Anzahl kodierter Textstellen je Ebene (gesamte Erhebung; n = 1436)
Ebene Departement / Abteilung	16	322	325
Ebene Institute / Labore / Forschungsgruppen	59	226	229
Ebene Professur	242	840	882

Aus 18 Evaluationsverfahren lagen Expertenberichte vor, die sich auf 20 Dokumente verteilten. Die Experten bezogen sich in ihren Evaluationsberichten auf 105 Institute, Labore und Forschungsgruppen sowie auf 201 Professuren. Insgesamt wurden 1 166 Textstellen in den Expertenberichten kodiert. Bei den Stellungnahmen der Departemente und der Hochschulleitung wurde bei der Auswertung auf die Unterscheidung nach verschiedenen Ebenen verzichtet, da die Dokumente diese Struktur nicht aufweisen. Die Ergebnisdarstellung konzentriert sich auf die über diese Analyseeinheit aggregierten Werte.

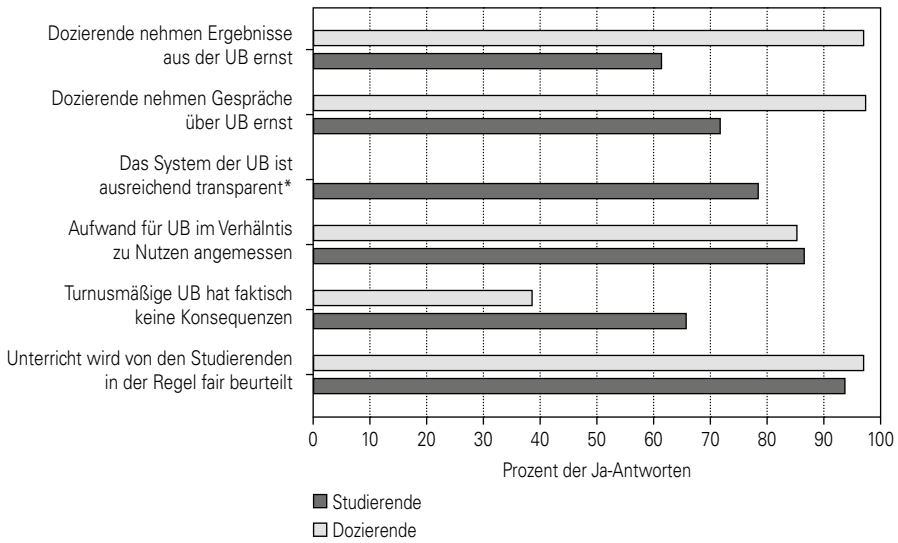
4 Zentrale Ergebnisse

4.1 Lehrveranstaltungsbeurteilung

Sowohl bei den Studierenden als auch bei den Dozierenden zeichnet sich insgesamt eine große Akzeptanz des gegenwärtigen Verfahrens der Lehrveranstaltungsbeurteilung ab. Studierende und Dozierende stimmen weitgehend und in hohem Maße (über 80 % bis 90 %) darin überein, dass Studierende Lehrveranstaltungen weitgehend fair beurteilen und dass das Verhältnis von Aufwand und Nutzen der Lehrveranstaltungsbeurteilung angemessen sei.

Während jedoch jeweils über 90 Prozent der Dozierenden angeben, die Ergebnisse der Lehrveranstaltungsbeurteilung und die Gespräche darüber mit den Studierenden ernst zu nehmen, wird diese Wahrnehmung nur von etwa 60 Prozent bzw. etwa 70 Prozent der Studierenden geteilt. Ferner schätzen sowohl die Dozierenden als auch die Studierenden die Frage nach den Konsequenzen der Lehrveranstaltungsbeurteilung kritisch ein: 39 Prozent der Dozierenden und 65 Prozent der Studierenden meinen, die Lehrveranstaltungsbeurteilung habe faktisch keine Konsequenzen (siehe Abbildung 1).

Abbildung 1: Bewertung der Lehrveranstaltungsbeurteilung (UB) durch die Studierenden und Dozierenden der ETH Zürich (Skalenwerte 4, 5 und 6 der 6-stufigen Likertskala – von 1 = „trifft überhaupt nicht zu“ bis 6 = „trifft voll und ganz zu“ wurden zu „Ja“-Antworten zusammengefasst; Studierenden- und Dozierendenbefragung, $N_{\text{Stud}}=2\,456$; $N_{\text{Doz}}=431$)



*Diese Aussage wurde in der Dozierendenbefragung nicht vorgegeben.

Die Frage, ob die Dozierenden Feedback-Gespräche mit den Studierenden über die Ergebnisse der Lehrveranstaltungsbeurteilung führen, bejahen 84 Prozent der Dozierenden und nur 68 Prozent der Studierenden. 16 Prozent der Dozierenden geben an, keine Gespräche zu führen und damit gegen die Weisung des Rektors zu verstoßen. Insbesondere von den Studierenden werden die mangelnden Konsequenzen der Lehrveranstaltungsbeurteilung angemahnt. Daher ist es in Zukunft erforderlich, transparenter zu machen, was mit den Ergebnissen der Lehrveranstaltungsbeurteilung passiert, welche Folgerungen auf den verschiedenen Ebenen für die Verbesserung der Lehre gezogen werden und dass die Befragungen auch tatsächlich Konsequenzen nach sich ziehen. 85 Prozent der befragten Professoren, die die Lehrveranstaltungsbeurteilung kennen (95% von 142 Professoren), meinen, dass sie sich bewährt habe (Werte 4–6 auf der 6-stufigen Antwortskala von 1 = „überhaupt nicht bewährt“ bis 6 = „voll und ganz bewährt“⁸; 59% vergeben die Werte 5–6; der Mittelwert beträgt 4.5).

Die Fragebögen zur Lehrveranstaltungsbeurteilung entsprechen zwar weitgehend den Richtlinien des Rektors, aber nicht ganz dem Stand des Wissens im Bereich der Lehr-

⁸ Die Antwortskala orientiert sich am Schweizer Schulnotensystem, bei dem die Note 1 für die schlechteste und die Note 6 für die beste Leistung steht.

evaluationsforschung. Auf die Einhaltung der folgenden Qualitätsstandards sollte geachtet werden:

1. Die Fragebögen sollten potentielle Bias-Variablen (z. B. Interesse am Thema unabhängig von der Darbietung, Wahl- oder Pflichtfach, Studienstufe) enthalten.
2. Studierende sollten die Möglichkeit haben, eine Gesamtbeurteilung der Lehrveranstaltung bzw. des Lehrenden vorzunehmen (z. B. „Ich bin mit der Lehrveranstaltung insgesamt sehr zufrieden“).
3. Die Fragebögen sollten weiterhin offene Fragen enthalten und den Dozierenden die Möglichkeit bieten, eigene Items zu formulieren.
4. Ergebnisse von Lehrveranstaltungsbeurteilungen, an denen weniger als 10 Personen teilgenommen haben, sollten nicht veröffentlicht werden. Ebenso sollten einzelne Fragen des Bogens, die weniger als 10 Personen unabhängig von der Teilnehmerzahl der Lehrveranstaltung beantworten, weder ausgewertet noch veröffentlicht werden.
5. Die Einwilligung der Dozenten zur Veröffentlichung der Evaluationsergebnisse sollte zu einem Zeitpunkt erbeten werden, zu dem die Dozierenden Kenntnisse über ihre Evaluationsergebnisse haben (so genannte informierte Einverständniserklärung).
6. Das Feedback für die Lehrenden sollte so aufgearbeitet sein, dass die Ergebnisse direkt auf Folie für einen Overheadprojektor kopierbar oder in eine Powerpointpräsentation integrierbar sind.
7. Es sollten weiterhin aus den Lehrveranstaltungsbeurteilungen je Departement/Studiengang und differenziert nach Typ der Lehrveranstaltung statistische Normen entwickelt werden, die den Dozierenden gestatten, die Ergebnisse ihrer Lehrveranstaltungsbeurteilung im Vergleich zu allen Lehrveranstaltungen des gleichen Typs eines Departements/Studiengangs zu beurteilen.
8. Für die Rückmeldung an die Dozierenden wäre es wünschenswert, auch bivariate Analysen zur Verfügung zu stellen, um beispielsweise in Erfahrung zu bringen, welche Aspekte der Lehre mit der Zufriedenheit der Studierenden oder mit dem selbst eingeschätzten Lernerfolg in Zusammenhang stehen.

Die testtheoretische Güte der Fragebögen im Hinblick auf die Reliabilität entspricht zwar den Erwartungen aus der wissenschaftlichen Literatur, sie ist aber relativ niedrig. Im Fragebogen erfasste Faktoren, die Einfluss auf das studentische Urteil haben, aber nicht mit der didaktischen Gestaltung der Lehre in Zusammenhang stehen (so genannte Bias-Variablen, z. B. Geschlecht), ist sehr schwach. Die Rücklaufquote der

Lehrveranstaltungsbeurteilung mit einem Median von 73 Prozent ist recht hoch und repräsentativ für die Population der Studierenden (Geschlecht und Muttersprache).

Angesichts der Vielzahl von eingesetzten Fragebögen, die sich inhaltlich zum Teil überlappen, sollte versucht werden, ETH-weit Bögen der Lehrveranstaltungsbeurteilung zu entwickeln, die den wissenschaftlichen Standards von Fragebögen zur Lehr-evaluation genügen und die die testtheoretischen Gütekriterien erfüllen (Reliabilität, Validität, Fairness). Verschiedene Departemente arbeiten bereits zusammen. Es sollte sowohl auf der Departementsebene wie auf der ETH-Ebene transparent sein (Homepage der ETH), welche Bögen zu welchem Zweck (Lehrveranstaltung, Seminare, Praktika, Assistierende) tatsächlich für die Lehrveranstaltungsbeurteilung eingesetzt werden.

4.2 Absolventenbefragung („Diplomiertenbefragung“)

85 Prozent der befragten Professoren, die die Absolventenbefragung kennen (64 % von 138 Professoren), meinen, dass sie sich bewährt habe (Werte 4–6 auf der 6-stufigen Antwortskala von 1 = „überhaupt nicht bewährt“ bis 6 = „voll und ganz bewährt“; 52 Prozent vergeben die Werte 5–6; der Mittelwert beträgt 4.5).

Die Absolventenbefragung genießt somit, sicherlich auch vor dem Hintergrund der wissenschaftlichen Fundiertheit der Befragung, eine hohe Akzeptanz bei den Professoren. Die Rücklaufquote der Absolventenbefragung liegt mit durchschnittlich 47 Prozent (Minimum: 43 %, Maximum: 56 %) in dem Bereich, der bei Absolventenstudien zu erwarten ist. Doch auch an der ETH Zürich besteht diesbezüglich Optimierungspotential.

Die Abstimmung der Absolventenbefragung mit weiteren Befragungen und ihre Einbindung in die Gesamtstrategie der ETH Zürich könnten optimiert werden: Absolventen der ETH Zürich werden zum einen im Rahmen der Beschäftigungsstatistik direkt nach Studienabschluss zur beruflichen Situation befragt, zum anderen im Rahmen der Departementsevaluation zur aktuellen beruflichen Situation sowie rückblickend zum Studium. Während die Beschäftigungsstatistik ETH-weit ausgewertet und die Ergebnisse Vertretern der Wirtschaft zugesendet werden, wird die Absolventenbefragung nur für das jeweilige evaluierte Departement ausgewertet und unter anderem den Experten im Rahmen des Peer Review zur Verfügung gestellt. Übergreifende Auswertungen erfolgten bisher nur im Rahmen dieser und der Vorläuferstudie. Daher wäre erstens zu empfehlen, die beiden Erhebungen im Rahmen einer Wiederholungsbefragung aufeinander zu beziehen: eine Befragung ein Jahr nach dem Studium rückblickend zu Lehre und Studium, und eine Wiederholungsbefragung zwei bis drei Jahre nach dem Studienabschluss eher zur Erwerbsbiographie und zum beruflichen Erfolg. Codes

könnten die beiden Fragebögen für eine Person identifizierbar machen, was unter anderem den Umfang der Fragen reduzieren würde. Zweitens wäre zu empfehlen, stärker mit dem Bundesamt für Statistik zusammenzuarbeiten, das regelmäßig Panel-Befragungen von Absolventinnen und Absolventen durchführt. Ergänzende Fragen der ETH Zürich könnten dem Fragebogen des Bundesamts für Statistik beigelegt werden. Die Beschäftigungsstatistik müsste dann im Hinblick auf den Befragungszeitpunkt angepasst, die Absolventenbefragung von der Evaluation eines Departements gegebenenfalls entkoppelt werden.

Entsprechend den dargestellten Befunden und der notwendigen Anpassung an die gestuften Studiengänge sollte eine Überarbeitung des Fragebogens der Absolventenbefragung, gegebenenfalls auch des Bogens der Beschäftigungsstatistik, vorgenommen werden (z. B. in Bezug auf Kompetenzen und Berufserfolg). So könnten Fragen aus nationalen oder internationalen Absolventenstudien (z. B. CHEERS, Kassel) übernommen werden, die einen Vergleichsrahmen für die Interpretation lokaler Ergebnisse an der ETH Zürich (z. B. Zeitdauer nach dem Studium bis zur ersten Erwerbstätigkeit) abgeben könnten.

Der Fragebogen der Absolventenbefragung orientiert sich zwar an wissenschaftlichen Standards, ist aber in verschiedenen Punkten verbesserungsbedürftig. So müsste der berufliche Erfolg als ein zentrales Kriterium eines erfolgreichen Studienabschlusses differenzierter erfasst werden: Es fehlen Fragen zur beruflichen Zufriedenheit. Es wird richtigerweise auf Kompetenzen statt auf Qualifikationen fokussiert, doch sind Begriffe wie Rahmenkompetenz, Fachwissen usw. unscharf formuliert und nicht deckungsgleich mit der in der wissenschaftlichen Literatur diskutierten Unterscheidung zwischen Fachkompetenz und Schlüsselkompetenz. Zudem fehlen unter anderem Fragen zur allgemeinen Studienzufriedenheit, zum Berufsverlauf, zur Abschlussnote und zur Studiendauer. Die rückblickenden Bewertungen von Lehrveranstaltungen und Dozierenden am Schluss des Fragebogens können stark durch zufällige Erinnerungseffekte beeinflusst werden, die nicht unbedingt im Zusammenhang mit der Qualität einer Lehrveranstaltung stehen.

Zudem sollten insbesondere die Funktionen dieser Befragung stärker herausgearbeitet werden: Neben den bisherigen Zielsetzungen der Absolventenbefragung, eine rückblickende Bewertung des Studiums zu erhalten oder über die aktuelle berufliche Situation der Absolventen zu informieren, könnte die Absolventenbefragung auch für die Alumni-Pflege eingesetzt werden (Erhebung des Weiterbildungsbedarfs, Adressenverwaltung, Informationen über die ETH, Stiftungen etc.). Ferner könnte überlegt werden, aus der Absolventenbefragung Indikatoren abzuleiten, die für die Hochschulleitung zur Steuerung von Lehre und Studium eingesetzt werden können (z. B. Berufserfolg, Studienzufriedenheit). Bezüglich der Auswertung der Ergebnisse sollte in Be-

tracht gezogen werden, ergänzend zur Grundauszählung multivariat-statistische Analysen zu rechnen, um beispielsweise Determinanten des beruflichen Erfolgs oder der Studiendauer zu bestimmen (z. B. Einfluss der Erwerbstätigkeit während des Studiums auf die Studiendauer und den Berufserfolg).

Neben der Untersuchung der Güte der Qualitätssicherungsinstrumente, welche im Vordergrund dieser Studie stand, wurden auch Ergebnisse zur Bewertung der Qualität der Lehre einbezogen. So weist die Absolventenbefragung nach, dass die Studieninhalte unabhängig von der Departementszugehörigkeit von 89 Prozent der Befragten ($n = 1\ 651$) als aktuell bezeichnet werden. Die ETH Zürich ist aus Sicht der Absolventen stark in der Vermittlung von Fachwissen und methodischem Wissen. Bezüglich der Vermittlung von Kenntnissen in Projekt- und Zeitmanagement bestehen Unterschiede zwischen den Departementen. Ehemalige Studierende erwarten von der ETH Zürich den Erwerb von Kompetenzen, die mit der wissenschaftlichen Tätigkeit zusammenhängen, weniger den Erwerb von überfachlichen Kompetenzen (z. B. Fremdsprachenkenntnisse, Wertschätzung von Kunst und Musik). Aus der Sicht der ehemaligen Studierenden wurden diese Erwartungen aber nur teilweise erfüllt. Am größten fallen die Unterschiede zwischen Erwartungen und Beiträgen der ETH Zürich bei den Kompetenzbereichen „Führungsqualität“, „Auftreten vor Publikum“, „Teamarbeit“ und „Kreativität“ auseinander. Insgesamt jedoch würde eine überwältigende Mehrheit von über 97 Prozent der Befragten wieder ein Studium ergreifen, und fast ein ebenso großer Prozentsatz von 93 Prozent würde wieder an der ETH Zürich studieren.

4.3 Selbstevaluation und Peer Review im Rahmen der Departementsevaluation

Die Analyse von Selbstevaluation und Peer Review hat gezeigt, dass etwa drei Viertel der 145 schriftlich befragten Professoren davon überzeugt sind, dass sich die Selbstevaluation bewährt hat (Mittelwert: 4.2); 69 Prozent teilen diese Einschätzung bezüglich des Peer Review (Mittelwert: 4.1; jeweils Werte 4–6 auf der 6-stufigen Antwortskala von 1 = „überhaupt nicht bewährt“ bis 6 = „voll und ganz bewährt“). Bei einer strengeren Dateninterpretation zeigt sich allerdings, dass jeweils weniger als die Hälfte der Befragten die Werte 5 bis 6 vergeben hat (Selbstevaluation: 44 %, Peer Review: 45 %).

Die Dokumentenanalyse hat gezeigt, dass in fast allen Selbstberichten (Dokumentationen) der Departemente und in den Evaluationsberichten der Experten insbesondere auf Departementsebene, aber auch auf Ebene der einzelnen Forschungsgruppen, Institute und Labore sowie der einzelnen Professuren Lehre und Studium thematisiert werden. In den Berichten werden jedoch auf allen drei Ebenen vor allem Texte rein deskriptiver Art formuliert (insbesondere Beschreibung des Lehrangebots bzw. des

Curriculums und der Lehrziele; Aufführen von Statistiken); dieser Anteil beträgt zwischen 73 Prozent (Departementsebene) und 84 Prozent (Ebene Institute bzw. Forschungsgruppen) aller Textstellen, die sich mit Lehre und Studium befassen. In den Evaluationsberichten der Experten liegt der deskriptive Anteil zwischen 37 Prozent (Departementsebene) und 61 Prozent (Ebene Professur). Die Empfehlungen im Bereich Lehre und Studium spielen hier eine größere Rolle, wobei die Experten ihre Empfehlungen insbesondere an die Departemente sowie an Institute, Abteilungen bzw. Labore und weniger an die einzelnen Professuren adressieren.

Dementsprechend nehmen bei der Thematisierung von Lehre und Studium in den *Dokumentationen* Aspekte wie Stärken und Schwächen, konkrete Zielsetzungen in der Lehre sowie innovative Lehrprojekte nur wenig Raum ein. So macht der Anteil aller Textstellen, die sich in den Dokumentationen mit Lehre und Studium befassen, zu dargestellten Stärken nur zwischen 6 Prozent (Ebene Professuren) und 9 Prozent (Ebene Departemente) aus. Der Anteil, in dem Schwächen benannt werden, umfasst auf allen drei Ebenen der Dokumentationen jeweils nur 6 Prozent der Textstellen. Der Anteil der Textstellen zu konkreten Zielsetzungen in Lehre und Studium (z. B. die Entwicklung eines interdisziplinären und hochschulübergreifenden Masterstudiengangs) beträgt nur 7 Prozent und zu den innovativen Lehrprojekten (z. B. Durchführung eines Pilotprojekts für die stärkere Beratung und Betreuung von Studierenden) lediglich 3 Prozent.

Die Ergebnisse von Lehrveranstaltungsbeurteilungen und Absolventenbefragungen spielen in den Dokumentationen sowie in den Evaluationsberichten der Experten keine oder nur eine geringe Rolle. Die Experten sprechen in ihren *Evaluationsberichten* Stärken in Lehre und Studium eher in Bezug auf die einzelnen Professuren an (etwa ein Fünftel aller Textstellen zu Lehre und Studium), weniger in Bezug auf die Departemente (5 Prozent aller Textstellen). Schwächen richten sich dagegen vor allem an das Departement (14 Prozent der Textstellen), weniger an die Professuren (3 Prozent aller Textstellen) (vgl. Tabelle 3). Der Konkretionsgrad der Empfehlungen ist auf Departementsebene am größten, er nimmt auf Institutsebene ab und fällt auf der Professurebene am geringsten aus. Dass sich die Kritik der Peers weniger an ihre zu evaluierende Peers und damit an (handlungsfähige) Individuen, sondern an eine übergeordnete Organisationseinheit wendet, bestätigt die häufig am Peer Review geäußerte Kritik, dass mit diesem Verfahren Bestehendes nicht stärker in Frage gestellt und damit auch keine größeren Veränderungen herbei geführt werden (eine Zusammenfassung zu den Stärken und Schwächen des Peer Review gibt unter anderem Bornmann 2011).

Tabelle 3: Inhalte im Bereich Lehre und Studium auf Departements-, Instituts- und Professurebene für 10 Evaluationsberichte der Experten; absteigend sortiert nach Departementsebene (erste Erhebungswelle: 10 Departemente/Abteilungen, 43 Institute und 59 Professuren, bei denen Lehre und Studium thematisiert werden; insgesamt 793 Textstellen)

Thema	Departement (n=365 Textstellen)		Institut (n=245 Textstellen)		Professur (n=181 Textstellen)	
	absolut	in Prozent	absolut	in Prozent	absolut	in Prozent
Deskriptive Inhalte	135	37 %	131	40 %	109	60 %
Empfehlungen	137	37 %	55	22 %	12	7 %
Schwächen	52	14 %	30	12 %	5	3 %
Stärken	18	5 %	27	11 %	38	21 %
Fragen der Hochschul- leitung	12	3 %	0	0 %	0	0 %
Lehrveranstaltungs- beurteilung	6	2 %	2	1 %	16	9 %
Bewertung von Qualitäts- sicherungsinstrumenten	2	1 %	0	0 %	0	0 %
Diplomiertenbefragung	3	1 %	0	0 %	1	1 %

In fast allen Stellungnahmen der Departemente und der Hochschulleitung werden Lehre und Studium thematisiert. Die Hochschulleitung spricht in ihren Stellungnahmen, die sich an den ETH-Rat richten, insbesondere die Stärken eines Departements in Lehre und Studium sowie die bereits ergriffenen Maßnahmen an (zusammen 60 Prozent aller Textstellen zu Lehre und Studium). Dieser Anteil macht in den Stellungnahmen der Departemente 38 Prozent der Textstellen aus. Die Departemente berichten stärker als die Hochschulleitung von geplanten bzw. beabsichtigten Maßnahmen (zusammen etwa ein Viertel der Textstellen) und pflichten den gutachterlichen Empfehlungen häufiger in allgemeiner Form bei (9 Prozent der Textstellen).

Ferner zeigt sich, dass in den Stellungnahmen der Departemente durchschnittlich insgesamt 20 Maßnahmen aufgeführt werden, die in Folge der Evaluation ergriffen wurden, geplant oder beabsichtigt sind. Von diesen 20 beziehen sich durchschnittlich sieben auf Lehre und Studium. Dabei wird durchschnittlich von drei bereits ergriffenen Maßnahmen berichtet, von zwei konkret geplanten sowie von zwei beabsichtigten Maßnahmen.

Vor dem Hintergrund der dargestellten Ergebnisse insgesamt könnte diskutiert werden, ob in den Evaluationsverfahren dem Bereich Lehre und Studium eine größere Bedeutung als bisher beigemessen und eine stärkere Akzentuierung auf Stärken, Schwächen, Probleme usw. vorgenommen werden soll. Der Stellenwert von Lehre

und Studium in den Evaluationsverfahren im Vergleich zu Forschung und Dienstleistung könnte für alle Verfahrensbeteiligten explizit gemacht werden. In diesem Zusammenhang könnte die Hochschulleitung auch prüfen, ob insbesondere auf der Ebene der einzelnen Professuren (sowohl bei der Selbstdarstellung als auch der Bewertung) im Verhältnis zur Departementsebene stärker als bisher auf Lehre und Studium eingegangen werden soll.

Ferner könnte geprüft werden,

1. inwieweit in den Selbstevaluationen und im Peer Review die Ergebnisse der Lehrveranstaltungsbeurteilungen und der Absolventenbefragungen (stärker) aufgenommen und kommentiert werden.⁹ So könnte z. B. ein Leitfaden, der den Departementen für die Erstellung ihrer Dokumentation im Rahmen der Selbstevaluation zur Verfügung gestellt wird (siehe nächster Punkt), den Hinweis enthalten, dass die Ergebnisse der Befragungen kommentiert werden sollen. Auch die Frageleitfäden der Hochschulleitung für die Erstellung des Berichts der Experten könnten entsprechend angepasst werden.
2. ob die Departemente von der Hochschulleitung einen detaillierten Leitfaden erhalten, der unter anderem die für die Beurteilung der Qualität von Lehre und Studium relevanten Aspekte klar benennt. Der Einsatz solcher Leitfäden gehört in der internationalen Evaluationspraxis zum Standard.¹⁰ Es sollte überlegt werden, ob in Zukunft bestimmte Aspekte von Lehre und Studium stärker in den Vordergrund gestellt werden (z. B. Beschreibung von Stärken, Schwächen, innovativen Lehrprojekten oder auch Lehrportfolios, didaktischen Qualifikationen oder auch Konzepten für Mobilität, Karriereperspektiven etc.).
3. ob für den Frageleitfaden für die Experten ein einheitliches Fragenset definiert werden soll, der jeweils departementsspezifisch ergänzt wird. Dies könnte auch zu einem einheitlicheren Aufbau der gutachterlichen Berichte führen (z. B. gemäß 2.4 der Standards and Guidelines von ENQA). Es könnte zudem in Erwägung gezogen werden, die Selbstberichte sowie die Evaluationsberichte der Experten in jedem Evaluationsverfahren dahingehend zu überprüfen, ob auf alle wichtigen Aspekte aus den Leitfäden der Hochschulleitung jeweils Bezug genommen wird bzw. ob gegebenenfalls Ergänzungen notwendig sind.

Die Phase der Umsetzung von Evaluationsergebnissen (Follow-Up bzw. Monitoring) gehört zu den zentralen Bestandteilen von mehrstufigen Evaluationsverfahren mit Peer

⁹ Es wäre sinnvoll, die verschiedenen Qualitätssicherungsinstrumente zielgerichtet und aufeinander bezogen einzusetzen (Mittag/Daniel 2008).

¹⁰ In den Standards and Guidelines von ENQA (European Association for Quality Assurance in Higher Education 2009) sind die Themen solcher Leitfäden formuliert.

Review (siehe z. B. Mittag/Bornmann/Daniel 2003; Mittag 2006). An der ETH Zürich ist diese bisher jedoch noch wenig institutionalisiert. Daher erscheint beispielsweise die Verabschiedung eines Maßnahmenprogramms sowie die Überprüfung von dessen Umsetzung in Form eines schriftlichen Umsetzungsberichts sinnvoll. Das Vorgehen zum Follow-Up bzw. Monitoring sollte klar definiert und einheitlich angewendet werden (z. B. gemäß 2.6 der Standards and Guidelines von ENQA).

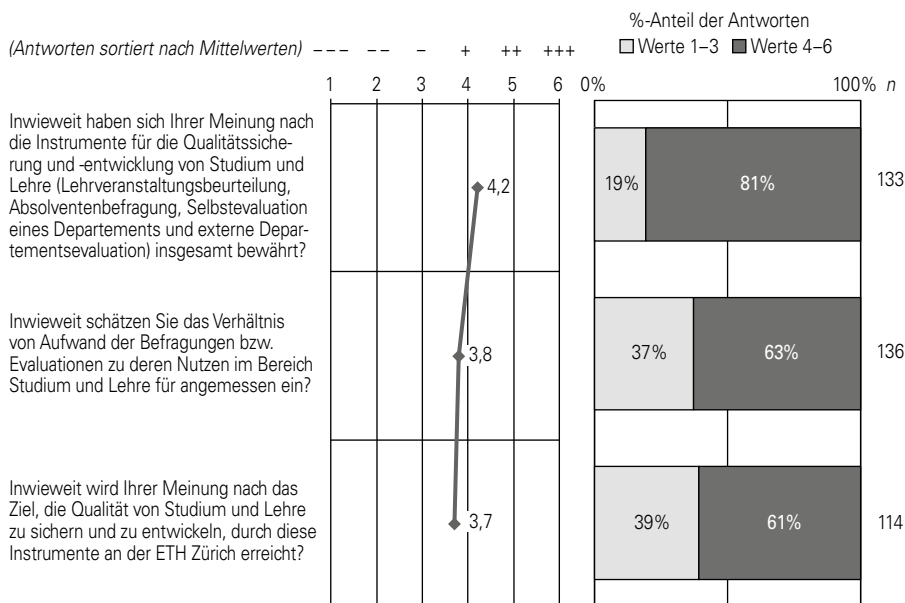
Das Verfahren sollte regelmäßig einer Meta-Evaluation unterzogen und die Einhaltung interner und externer Standards und Richtlinien überprüft werden (Richtlinien der Schweizerischen Universitätskonferenz vom 7. Dezember 2006; Qualitätsstandards des OAQ, Organ für Akkreditierung und Qualitätssicherung der Schweizerischen Hochschulen 2007; Standards and Guidelines von ENQA, European Association for Quality Assurance in Higher Education 2009; Standards der Schweizerischen Evaluationsgesellschaft SEVAL, Widmer/Landert/Bachmann 2000).

4.4 Gesamteinschätzung

Gefragt nach ihrer Gesamteinschätzung zu allen vier Instrumenten der Qualitätssicherung und -entwicklung (Lehrveranstaltungsbeurteilung, Absolventenbefragung sowie Selbstevaluation und Peer Review im Rahmen der Departementsevaluation) geben 81 Prozent der befragten Professoren an, dass sich die Instrumente insgesamt bewährt haben (Werte 4–6 auf der 6-stufigen Antwortskala von 1 = „überhaupt nicht bewährt“ bis 6 = „voll und ganz bewährt“), der Mittelwert beträgt 4.2 (Abbildung 2). Jeweils mehr als 60 Prozent der Befragten meinen, dass das Verhältnis von Aufwand für die Befragungen bzw. Evaluationen und deren Nutzen angemessen ist und dass das Ziel, die Qualität von Lehre und Studium an der ETH Zürich zu sichern und zu entwickeln, durch diese Instrumente erreicht wird. Die Mittelwerte fallen insbesondere bei der Frage nach dem Verhältnis von Aufwand und Nutzen und bei der Frage nach der Zielerreichung mit 3.8 bzw. 3.7 eher gering aus.¹¹

¹¹ Auch in der Studie von Mittag/Bornmann/Daniel (2003) schätzen die Fachangehörigen das Verhältnis von Aufwand und Ertrag eher kritisch ein.

Abbildung 2: Gesamtbewertungen zu den Qualitätssicherungsinstrumenten an der ETH Zürich (Antworten der Professoren; Mittelwert und Prozentanteile für zwei Antwortkategorien)



Vor dem Hintergrund dieser Ergebnisse sollte kritisch geprüft werden, ob für alle vier Qualitätssicherungsinstrumente (Selbstevaluation, Peer Review, Lehrveranstaltungsbeurteilung und Absolventenbefragung) im Bereich Lehre und Studium das Verhältnis von Aufwand und Nutzen noch optimiert und damit die Wirkung der Befragungen und Evaluationsverfahren in diesem Bereich weiter erhöht werden kann (siehe hierzu z. B. auch D3 und N8 der SEVAL-Standards).

Knapp zwei Drittel der Professoren sind der Auffassung, dass das System des Qualitätsmanagements an der ETH Zürich ausreichend transparent ist (Werte 4 bis 6 auf der sechsstufigen Antwortskala; 42 Prozent haben die Werte 5 bis 6 vergeben). Der Mittelwert beträgt lediglich 3.4. Daher sollte geprüft werden, ob das Qualitätsmanagementsystem für alle Beteiligten transparenter gestaltet werden kann (gemäß 3.1 der OAQ-Standards sowie 1.1 der Standards and Guidelines von ENQA).

5 Schlussfolgerungen

Die ETH Zürich verfügt über etablierte Qualitätssicherungsinstrumente im Bereich Lehre, die in der Hochschule weitestgehend akzeptiert sind. Während an manchen Universitäten lediglich ein einziges Instrument für die Qualitätssicherung der Lehre,

nämlich die Lehrveranstaltungsbeurteilung eingesetzt wird (z. B. Universität Innsbruck: Schafferer 2010), setzt die ETH Zürich auf ein „Evaluationssystem nach dem Baukastensystem“ (Handbuch zum Qualitätsmanagement, Niedermann/Hugentobler/Osterwalder 2003, S. 5).

Die Meta-Evaluation hat allerdings auch gezeigt, dass es noch eine Reihe von Verbesserungsoptionen gibt. So sollten insbesondere die verschiedenen Instrumente noch stärker aufeinander bezogen werden. Im Schlussbericht des Quality-Audits aus dem Jahr 2007/2008 wird dementsprechend festgehalten: „Insbesondere empfiehlt das OAQ, ein noch kohärenteres Qualitätssicherungssystem zu entwickeln. Mit Nachdruck unterstützt es die ETH Zürich in ihrem Bestreben, die Qualitätssicherung in der Lehre noch stärker zu systematisieren...“ (Organ für Akkreditierung und Qualitätssicherung 2008, S. 15).

Trotz der Einführung von Qualitätsmanagementsystemen für Studium und Lehre im Zusammenhang mit der Systemakkreditierung in Deutschland bestehen an vielen Hochschulen die verschiedenen Qualitätssicherungsinstrumente immer noch parallel nebeneinander. Die Einbindung der Instrumente in ein gesamtes Qualitätsmanagementsystem ermöglicht jedoch eine Effizienzsteigerung der Instrumente (Synergieeffekte und Vermeidung von Parallelarbeit) und liefert zielgerichtete Informationen für die Entscheidungsträger. Die verschiedenen Instrumente verfolgen unterschiedliche Zielsetzungen, die klar definiert sein sollten. Gleichzeitig muss ihr Nutzen bzw. die Umsetzung der jeweiligen Ergebnisse gewährleistet sein. Werden diese Voraussetzungen erfüllt, kann davon ausgegangen werden, dass die verschiedenen Qualitätssicherungsinstrumente bzw. auch das Qualitätsmanagementsystem als Ganzes bei den Hochschulangehörigen auf größere Akzeptanz als bisher stoßen. Die „Qualitätsbürokratie“, die vielerorts kritisiert wird, könnte in einem in sich abgestimmten Qualitätsmanagementsystem auf ein Mindestmaß reduziert werden. Die Entwicklung von Qualitätsmanagementsystemen bietet damit auch die Chance, bisher Versäumtes nachzuholen (siehe auch Mittag/Daniel 2008).

Darüber hinaus weist der Expertenbericht der Zwischenevaluation des ETH-Bereichs (vgl. Peer Review Committee 2011, S. 9–10) darauf hin, dass die eingesetzten Befragungsinstrumente noch nicht den Wechsel vom Lehrenden zum Lernenden vollzogen haben und dass Möglichkeiten gefunden werden sollten, um die Lernleistung der Studierenden zu messen. Seit der Durchführung des Audit ist einige Zeit vergangen, und die ETH Zürich hat bereits mit der Einleitung entsprechender Maßnahmen begonnen. Dieser Perspektivenwechsel wurde auch im Zuge des Bologna-Prozesses gefordert; viele Hochschulen befinden sich diesbezüglich in einem noch andauernden Wandlungsprozess. So besteht beispielsweise auch an einigen bayerischen Hochschulen noch Handlungsbedarf in diesem Bereich (Witte/Sandfuchs/Mittag/Brummerloh 2011).

Im Bereich der mehrstufigen Evaluationen mit Peer Review ist im Rahmen der Studie deutlich geworden, dass dieses Instrument für die Qualitätssicherung und -entwicklung von Lehre und Studium für die befragten Professoren einen geringeren Stellenwert einnimmt als die Lehrveranstaltungsbeurteilung und die Absolventenbefragung. Die Evaluationen könnten an der ETH Zürich somit noch viel stärker für die Qualitätssicherung und -entwicklung von Lehre und Studium eingesetzt bzw. genutzt werden. Auch an anderen Hochschulen gibt es mehrstufige Evaluationsverfahren mit Peer Review, die sowohl die Forschung als auch Lehre und Studium zum Gegenstand haben (z. B. die Universität Zürich, die Universität Mainz und die Universität Duisburg-Essen). Bei diesen Verfahren muss beiden Bereichen entsprochen werden: Das Ziel, die Qualität zu sichern und zu entwickeln, sollte für beide Bereiche erreicht werden. Sind Schwerpunktsetzungen gewollt und gewünscht, sollte dieses ausreichend innerhalb der Hochschule kommuniziert und transparent gemacht werden. An der ETH Zürich ist die Phase der Umsetzung von Evaluationsergebnissen (Follow-Up bzw. Monitoring) noch wenig institutionalisiert. Sie gehört jedoch zu den zentralen Bestandteilen von mehrstufigen Evaluationsverfahren mit Peer Review (siehe z. B. Mittag/Bornmann/Daniel 2003; Mittag 2006).

Zudem konnte durch die Metaanalyse festgestellt werden, dass auch die Lehrveranstaltungsbeurteilung und die Absolventenbefragung Optimierungspotentiale besitzen, zum einen die Anpassung an den aktuellen wissenschaftlichen Standard betreffend, zum anderen vor allem in Bezug auf einen abgestimmten und effektiveren Verfahrensablauf. Die ETH Zürich ist derzeit dabei, die beiden Befragungstypen zu überarbeiten. Die wissenschaftliche Auseinandersetzung mit beiden Befragungstypen ist weit fortgeschritten. Alle Hochschulen sollten sich bemühen, die (bereits definierten) Standards umzusetzen. Absolventenstudien geben wertvolle Hinweise unter anderem in Bezug auf die Bewertung ihres abgeschlossenen Studiums durch Ehemalige und sollten ein zentraler Bestandteil von Qualitätssicherungssystemen an Hochschulen sein. Die vorliegende Analyse zeigt, dass die rückblickende Bewertung des Studiums durch die Absolventen der ETH Zürich insgesamt gesehen sehr positiv ausfällt, wenn auch bestimmte Bereiche noch verbessert werden können.

Die ETH Zürich hat ihre Qualitätssicherungsinstrumente in Lehre und Studium einer umfassenden Meta-Evaluation unterzogen. Sie hat die Studie in Auftrag gegeben, umfangreiches Material zur Verfügung gestellt, und sie hat die Ergebnisse umfassend in das Quality-Audit des Organs für Akkreditierung und Qualitätssicherung der Schweizerischen Hochschulen 2007/2008 eingebracht. An diesem offensiven, selbstkritischen und optimierungsorientierten Ansatz sollten sich andere Hochschulen ein Beispiel nehmen – die Anzahl an Vergleichsstudien ist verschwindend gering.

Literatur

Abd-el-Razik, Adham; Eugster, Balthasar; Grund, Sven; Lutz, Leonard; Mittag, Sandra et al. (2007): Standortbestimmung durch die Analyse des Ist-Zustandes der Qualität von Lehre und Studium an der ETH Zürich (Projekt: „Roadmap: Maßnahmen zur Verbesserung von Lehre und Studium an der ETH Zürich“). Zürich

Alean-Kirkpatrick, Pamela; Hänni, Hans; Lutz, Leonard (1997): Internal quality monitoring of the teaching at the ETH Zurich: model design and initial impact. In: *Quality in Higher Education*, 3, 1997, 1, S. 63–71

Bornmann, Lutz (2011): Scientific peer review. In: *Annual Review of Information Science and Technology*, 45, 2011, S. 199–245

European Association of Quality Assurance in Higher Education (2009): Standards and Guidelines for Quality Assurance in the European Higher Education Area. Helsinki

Kübler, Olaf (2003): Lessons from a Decade of Peer Review at ETH Zürich. In: Max-Planck-Gesellschaft (Hrsg.): *Science between Evaluation and Innovation: A Conference on Peer Review*. München, S. 67–72

Mayring, Philipp (2008): Qualitative Inhaltsanalyse: Grundlagen und Techniken. Beltz

Mittag, Sandra (2006): Qualitätssicherung an Hochschulen. Eine Untersuchung zu den Folgen der Evaluation von Lehre und Studium (Internationale Hochschulschriften, Bd. 475, Reihe für Habilitationen und sehr gute und ausgezeichnete Dissertationen). Münster

Mittag, Sandra; Mutz, Rüdiger; Daniel, Hans-Dieter (2012): Institutionelle Qualitätssicherung der Lehre auf dem Prüfstand: Eine Fallstudie an der ETH Zürich. Bielefeld

Mittag, Sandra; Bornmann, Lutz; Daniel, Hans-Dieter (2003): Evaluation von Lehre und Studium an Hochschulen. Handbuch zur Durchführung mehrstufiger Evaluationsverfahren. Münster

Mittag, Sandra; Daniel, Hans-Dieter (2008): Qualitätsmanagement an Hochschulen (Vorabdruck). Qualität in der Wissenschaft. In: *Zeitschrift für Qualitätsentwicklung in Forschung, Studium und Administration (Systemakkreditierung und Qualitätsmanagement an Hochschulen)*. 1, 2008, 2, S. 13–18

Mutz, Rüdiger; Daniel, Hans-Dieter (2008): Kompetenzadäquanz der Beschäftigung von Hochschulabsolvent/innen – untersucht am Beispiel der ETH Zürich. In: B. M. Kehm (Hrsg.): *Hochschule im Wandel. Die Universität als Forschungsgegenstand*. Festschrift für Ulrich Teichler. Frankfurt am Main, S. 97–106

Niedermann, Christoph; Hugentobler, Urs; Osterwalder, Konrad (2003): Handbuch zum Qualitätsmanagement der ETH Zürich von Lehre, Forschung und Dienstleistungen an der ETH Zürich.

Organ für Akkreditierung und Qualitätssicherung (2007): Quality Audits 2007/08. Konzept, Ablauf und Qualitätsstandards. Leitfaden für die universitären Hochschulen. Bern

Organ für Akkreditierung und Qualitätssicherung (2008): Schlussbericht OAO, ETH Zürich, September 2008. http://www.oaq.ch/pub/de/documents/Schlussbericht_ETHZ.pdf (Zugriff: 17. Januar 2012)

Peer Review Committee (2011): Zwischenevaluation des ETH-Bereichs. Evaluation durch die Experten vom 7.-10. November 2010. Expertenbericht vom November 2010 (Übersetzung des englischen Originals durch den ETH-Rat vom 25.2.2011). www.sbf.admin.ch/htm/dokumentation/publikationen/uni/Zwischenevaluation_ETH_Bereichs_2010.pdf (Zugriff: 17. Januar 2012)

Schafferer, Wolfgang Gregor (2010): Lehrevaluation und Studiensituation an der Universität Innsbruck. Frankfurt am Main

Schweizerische Universitätskonferenz (2007): Richtlinien für die Qualitätssicherung an den schweizerischen universitären Hochschulen (Qualitätssicherungs-Richtlinien) vom 7. Dezember 2006. <http://www.cus.ch/wDeutsch/publikationen/richtlinien/D-443-06A-Quali-RL-VO.pdf> (Zugriff: 17. Januar 2012)

Ulich, Eberhard (1999): Qualitätsmanagement an Universitäten und Hochschulen. In: Thonhauser, Josef u.a. (Hrsg.): Evaluation im Bildungsbereich: Wissenschaft und Praxis im Dialog. Innsbruck, S. 85–102

Widmer, Thomas; Landert, Charles; Bachmann, Nicole (2000): Evaluations-Standards der Schweizerischen Evaluationsgesellschaft (SEVAL-Standards). http://www.seval.ch/de/documents/seval_Standards_2001_dt.pdf (Zugriff: 17. Januar 2012)

Witte, Johanna; Sandfuchs, Gabriele; Mittag, Sandra; Brummerloh, Sven (2011): Stand und Perspektiven bayerischer Bachelor- und Masterstudiengänge. Studien zur Hochschulforschung (82). München: IHF.

Anschriften der Verfasser:

Dr. Sandra Mittag
Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF)
Prinzregentenstraße 24
80538 München
E-Mail: mittag@ihf.bayern.de

Dr. Rüdiger Mutz
Prof. Dr. Hans-Dieter Daniel
ETH Zürich
Departement Geistes-, Sozial- und Staatswissenschaften
Professur für Sozialpsychologie und Hochschulforschung
Mühlegasse 21
CH- 8001 Zürich
E-Mail:
mutz@gess.ethz.ch
daniel@gess.ethz.ch

Dr. Sandra Mittag war 2007–2009 Postdoktorandin an der Professur für Sozialpsychologie und Hochschulforschung der ETH Zürich und arbeitete bis Juni 2012 als wissenschaftliche Referentin am Bayerischen Staatsinstitut für Hochschulforschung und Hochschulplanung. Seit Juli 2012 ist sie Geschäftsführerin bei der Stiftung „Gesellschaft macht Schule“.

Prof. Dr. Hans-Dieter Daniel ist Ordinarius und Dr. Rüdiger Mutz wissenschaftlicher Mitarbeiter an der Professur für Sozialpsychologie und Hochschulforschung der ETH Zürich.

Qualitätssicherung in Großbritannien durch die Quality Assurance Agency – Ein Überblick

Anett Löscher

Die Quality Assurance Agency for Higher Education (QAA) prüft, ob und wie britische Universitäten, university colleges und further education colleges die akademischen Standards ihrer Bildungsangebote wahren und sichern, und wie sie die Qualität ihres Lehr- und Lernangebots kontinuierlich weiterentwickeln. Für die vier Länder des Vereinigten Königreichs – England, Wales, Nordirland und Schottland – hat die QAA jeweils verschiedene Begutachtungsmethoden entwickelt, die sowohl den jeweiligen bildungspolitischen Kontexten der vier Nationen als auch den strukturellen und organisatorischen Unterschieden von Bildungsanbietern Rechnung tragen. Im Folgenden wird ein Überblick über die verschiedenen Begutachtungsmethoden gegeben, die die QAA entwickelt hat und anwendet. Auch wird der UK Quality Code for Higher Education (UKQCHE) vorgestellt, der Referenzpunkt ist für die Sicherung akademischer Standards und für die Weiterentwicklung der Qualität von Lernmöglichkeiten, und damit Grundlage für die Begutachtungen der QAA.

1 Einführung

Das britische Bildungswesen unterscheidet zwischen beruflicher (vocational education) und höherer Bildung (higher education). Der Begriff „higher education“ umschreibt alle Arten post-sekundärer akademischer Ausbildung, die zu akademischen Qualifikationen führt. „Höhere Bildung“ ist einerseits das Metier der britischen Universitäten. Im recht differenzierten britischen Hochschulsektor spielen jedoch andere Bildungseinrichtungen eine (zunehmend) wichtige Rolle: University colleges (Einrichtungen mit weniger als 4.000 Studierenden und oft sehr spezialisiertem Lehrangebot), die eher deutschen Berufsschulen vergleichbaren further education colleges sowie privat finanzierte for-profit und non-profit Einrichtungen bilden Studierende auf Universitätsniveau und mit anerkannten britischen akademischen Qualifikationen aus, wenngleich sie selbst (von wenigen Ausnahmen abgesehen) nicht das Recht haben, diese Abschlüsse zu vergeben. So gibt es z. B. in England 2012 89 Universitäten – die Gesamtziffer aller „Einrichtungen höherer Bildung“ (higher education institutions), die im tertiären Bereich akademische Qualifikationen vermitteln, liegt jedoch für England bei 131¹. Einrichtungen, die nicht das Recht haben, akademische Abschlüsse zu verleihen (sie haben in der britischen Terminologie keine *degree-awarding powers*) bieten ihr Lehrangebot in

¹ Vgl. <http://www.universitiesuk.ac.uk/UKHESector/Pages/OverviewSector.aspx> (Zugriff 10.01.2012)

Kooperation mit Universitäten oder Organisationen wie Edexcel an, die z. B. entweder Lehrangebote einer Einrichtung validieren oder eigene Studienprogramme als franchise an eine Einrichtung geben.

Alle Einrichtungen, die Studierende auf Universitätsniveau und zu anerkannten britischen akademischen Abschlüssen hin ausbilden, sind für die Qualität ihrer Lehrangebote verantwortlich und müssen über entsprechende interne Qualitätssicherungssysteme verfügen. Für die Wahrung der akademischen Standards ihrer Qualifikationen sind ultimativ die Einrichtungen verantwortlich, die das Recht haben, diese Abschlüsse zu vergeben – folgerichtig müssen z.B. Universitäten, deren Lehrangebot durch ein further education college vermittelt wird, nachweisen können, dass die akademischen Standards dieses Lehrangebots kontinuierlich gewahrt werden. Auch hierfür sind angemessene interne Qualitätssicherungssysteme zu etablieren. Bildungseinrichtungen sind also autonom in der Entwicklung ihrer Lehre und zugleich verantwortlich für deren Qualitätssicherung. Die Begutachtungen der QAA dienen mithin der externen Qualitätssicherung interner Strukturen und Prozesse.

2 Bildungspolitischer Kontext

Höhere Bildung ist Angelegenheit der vier Länder des Vereinigten Königreichs, wobei die Bildungspolitik Englands oft leitmotivisch wirkt. Das ist zum einen der Größe des englischen Sektors geschuldet, zum anderen der Dezentralisierung, die bewirkt, dass bildungspolitische Leitlinien der Londoner Zentralregierung oft direkt auf den englischen Sektor wirken, wohingegen die Definitions- und Gestaltungsmacht der walisischen und schottischen Regionalregierungen der entsprechenden Bildungspolitik zum Teil einen ganz anderen und sehr länderspezifischen Fokus geben. In ihren jährlichen Zuweisungen von Mitteln an den jeweiligen funding council identifizieren die Regionalregierungen zugleich Investitionsschwerpunkte sowie längerfristige, durch entsprechende Strategien zu erreichende Zielvorgaben. So besteht eine Zusammenarbeit zwischen Welsh Assembly Government und Higher Education Funding Council for Wales, zwischen Department for Business, Innovation and Skills und Higher Education Funding Council for England, sowie zwischen Scottish Parliament und Scottish Funding Council.

Der jeweils aktuelle bildungspolitische Kontext beeinflusst wesentlich die Arbeit der QAA, insbesondere die Entwicklung von Begutachtungsmethoden. Die Position der QAA als Auftragnehmer der funding councils, die ja ihrerseits durch ministerielle *grant letter* mit der Einführung, Fortsetzung und Implementierung von bildungspolitischen Maßnahmen beauftragt sind, begründet dies. Zudem sind Natur, Rolle, Zweck und Effektivität externer Qualitätssicherung seit dem Bericht eines parlamentarischen Komitees im Jahr 2009 (*Select Committee 2009*) aus verschiedenen Perspektiven hinterfragt worden und haben die QAA unter anderem veranlasst, wesentlich aktiver

die Öffentlichkeit zu suchen sowie über ihre Rolle und Arbeit Auskunft zu geben. In jüngster Zeit haben unter anderen die Umstrukturierung der staatlichen Hochschulfinanzierung zusammen mit der Einführung höherer Studiengebühren (*Browne 2010; BIS 2011a*),² die Änderung der Zuwanderungsbestimmungen für Studierende aus dem nicht-europäischen Ausland (*UKBA 2011*, siehe Fußnote 8), sowie die Ankündigung des Departments for Business, Innovation and Skills, die Regulierungslandschaft für höhere Bildung ändern zu wollen (*BIS 2012*), auf die Positionierung und Arbeit der QAA im britischen Hochschulsektor gewirkt.

3 Die Quality Assurance Agency for Higher Education (QAA)³

Die QAA entwickelt die Parameter und Methoden für die externe Qualitätssicherung höherer Bildung in ganz Großbritannien, ist also nicht nur für den englischen Sektor zuständig. Unter die Zuständigkeit der QAA fallen alle britischen staatlich finanzierten Universitäten, staatlich finanzierte university colleges, further education colleges und seit neuestem auch privat finanzierte for-profit und non-profit Einrichtungen, die Studierende auf Universitätsniveau und zu anerkannten britischen Abschlüssen hin ausbilden. Mithin prüft die QAA auch die Qualitätssicherungsstrukturen britischer Hochschulanbieter, die im Ausland operieren und Studierende dort zu britischen Hochschulabschlüssen führen.

Die QAA führt die Begutachtungen des Qualitätsmanagements von Bildungsanbietern als Auftragnehmer der britischen higher education funding councils durch. Der Further and Higher Education Act 1992 und der Further and Higher Education (Scotland) Act 1992 haben unter anderem die Etablierung von funding councils für Wales, Schottland und England beschlossen. Funding councils sind Behörden, die jährlich von der entsprechenden Regionaladministration Budgets für die Hochschulbildung in England, Schottland, Wales und Nordirland erhalten und diese dann an die jeweiligen, durch staatliche Gelder finanzierten Einrichtungen in den vier Ländern verteilen. Diese Verteilung erfolgt zum einen entsprechend der Maßgaben in den ministeriellen *grant letters*, die die Budgets begleiten und de facto Zielvorgaben für den jeweiligen nationalen Hochschulsektor formulieren, zum anderen entsprechend von Berechnungsmodellen,

² Hier ist wesentlich anzumerken, dass Studierende nicht pro Studienjahr eine finanzielle Transaktion in Höhe von bis zu 9.000 GBP an die ausbildende Hochschule vornehmen müssen. Vielmehr werden die Ausbildungskosten durch quasi den Hochschulen vorgeschossene öffentliche Gelder gedeckt. Studierende zahlen diese vorgetreckten Gelder erst nach Studienabschluss und nach erfolgreichem Eintritt ins Berufsleben in Form einer Graduiertensteuer zurück. Diese wird proportional zum Einkommen berechnet; das Mindesteinkommen, ab dem diese Steuer fällig wird, liegt bei ca. 21.000 GBP im Jahr. Problematisch ist dieses Finanzierungssystem unter anderem durch die signifikanten, zunächst aus dem Haushalt vorzustreckenden Summen, sowie durch die potentiell hohen Zahlen von Nicht- oder Teilrückzahlern (z. B. Absolventen, die die Einkommensgrenze nicht oder erst spät erreichen, oder aus dem britischen Steuersystem durch Beschäftigung im Ausland ausscheiden). Weitere Details sind abrufbar via <http://www.bis.gov.uk/studentfinance/>

³ Vgl. www.qaa.ac.uk

die unter anderem Größe, Profil und Studierendenzahlen der jeweiligen Hochschuleinrichtung berücksichtigen. Die funding councils sind in ihrer Eigenschaft als Verwalter des Budgets für Hochschulbildung gesetzlich verpflichtet, die Qualität der Bildung, die sie ja im öffentlichen Auftrag finanzieren, zu prüfen, zu wahren und wo immer möglich innovativ voranzutreiben. Diesen Auftrag haben die funding councils an die QAA übertragen und finanzieren die QAA entsprechend. Des Weiteren sind alle öffentlich finanzierten Hochschulen, die von der QAA begutachtet werden, de facto Abonnenten der Dienstleistungen der QAA und entrichten einen jährlichen Beitrag an die QAA. Unter anderem zum Zweck der Reputationspflege haben sich auch einige Einrichtungen aus dem privat finanzierten Bildungsbereich für ein Abonnement bei der QAA entschieden.

4 Referenzpunkt für die Qualitätssicherung – der UK Quality Code for Higher Education⁴

Die QAA entwickelt und pflegt im Auftrag und in Kooperation mit dem britischen Hochschulsektor den UK Quality Code for Higher Education (UKQCHE) als definitiven nationalen Referenzpunkt für die Qualitätssicherung. Die individuellen Teile und Kapitel der UKQCHE werden jeweils zusammen mit Arbeitsgruppen entwickelt, die zum einen repräsentativ für die Diversität der Bildungseinrichtungen im britischen Hochschulsektor sind, und zum anderen die spezifische Expertise versammeln, die zur Gestaltung des jeweiligen Kapitels notwendig ist. Alle Teile des UKQCHE durchlaufen eine Phase öffentlicher Konsultation und werden nach abschliessender Redaktion auf den Internetseiten der QAA veröffentlicht. Der UKQCHE gilt für alle Bildungseinrichtungen, deren Bildungsprogramme entweder zu einem britischen akademischen Abschluss führen oder deren Programme von einer britischen Einrichtung mit „*degree-awarding powers*“ validiert wurden. Er findet Anwendung in allen vier Verwaltungen des Vereinigten Königreichs (Schottland, England, Wales, Nordirland) und gilt zudem für alle britischen Hochschulprogramme, die in Übersee angeboten werden. Der Quality Code wahrt mithin die Interessen aller, die unter diese Definition britischer Hochschulbildung fallen.

Der UKQCHE legt Anforderungen (*Expectations*) fest, die der Hochschulsektor gemeinschaftlich und im Sinne der Selbstregulierung als Mindestanforderungen an sich selbst beschlossen hat. Als Referenzpunkt für die interne und externe Qualitätssicherung soll der UKQCHE somit akademische Standards britischer Hochschulbildung schützen und wahren, die Qualität der Lernmöglichkeiten sichern, die Hochschuleinrichtungen ihren Studierende verfügbar machen, die kontinuierliche und systematische Qualitätsentwicklung britischer Hochschulbildung unterstützen und vorantreiben, sowie

⁴ Siehe im Detail <http://www.qaa.ac.uk/AssuringStandardsAndQuality/quality-code/Pages/default.aspx>, Zugriff 10.01.2012

die Zugänglichkeit und Qualität von Informationen, die britische Hochschuleinrichtungen über sich selbst verbreiten, für die breite Öffentlichkeit sichern.

4.1 Akademische Standards

Der UKQCHE definiert akademische Mindeststandards als *“the minimum acceptable level of achievement that a student has to demonstrate to be eligible for an academic award”* (QAA 2011f, S. 11). „Award“ meint dabei die Zuerkennung sowohl akademischer Qualifikationen und Grade (z. B. Foundation Degree oder Bachelor) als auch von credits bei erfolgreicher Absolvierung eines Teils eines Studienprogramms durch Studierende. Für einander äquivalente Zuerkennungen von Qualifikationen oder credits sollten Studierende also gleiche Mindestleistungen erbringen müssen, ungeachtet dessen, durch welche britische Hochschuleinrichtung die jeweilige Zuerkennung erfolgt. Hochschuleinrichtungen sind darüber hinaus dafür verantwortlich, Kriterien festzulegen, mittels derer die Qualität von Studienleistungen über die Mindestanforderung für einen bestimmten award hinaus bewertet und differenziert werden können – anhand welcher Kriterien wird z. B. eine Seminararbeit als „gut“ bewertet oder ein Bachelor zusammenfassend als „mit Auszeichnung“ abgeschlossen klassifiziert?

Teil A des UKQCHE definiert die Referenzpunkte, anhand derer Bildungsanbieter akademische Standards setzen und pflegen: das *Framework for higher education qualifications in England, Wales and Northern Ireland* (FHEQ) und das *Framework for qualifications of higher education institutions in Scotland* (FQHEIS), die *Subject Benchmark Statements* und die *Foundation Degree qualification benchmark* (FDQB), sowie Richtlinien für Entwicklung, Validierung, Management und Revision von Studiengängen und ihren konstituierenden Einheiten, Richtlinien für die Nutzung externer Prozesse und Strukturen zum Zweck der kontinuierlichen Qualitätssicherung und -entwicklung, Richtlinien für das Monitoring von Lernergebnissen.

4.2 Qualität und Qualitätsentwicklung (“Enhancement”)

Der UKQCHE beschreibt Qualität als *“how well the learning opportunities made available to students enable them to achieve their award. It is about making sure that appropriate and effective teaching, support, assessment and learning resources are provided for them”* (QAA 2011f, S. 11). Die QAA sieht die Verantwortung für die Existenz und Qualität von Lernmöglichkeiten klar bei der Hochschuleinrichtung. Diese muss sicherstellen, dass Regularien, Strukturen und Prozesse für das Management aller Lernmöglichkeiten effizient etabliert und genutzt werden, so dass die Lernmöglichkeiten selbst von größtmöglicher Qualität und Nutzen sind. Einer Hochschuleinrichtung obliegt jedoch keine Verantwortung dafür, dass und wie diese Lernmöglichkeiten von Studierenden genutzt und erfahren werden.

Jede Begutachtung durch die QAA hat nicht nur die Sicherung von Standards und Qualität zum Ziel, sondern auch die Identifizierung guter Praxis und Anregung zur kontinuierlichen innovativen Weiterentwicklung der Qualität von Bildungsangeboten und Lernmöglichkeiten. Dies meint der UKQCHE (und folgerichtig die Begutachtungsmethoden der QAA) mit dem Begriff „*enhancement*“. Im Gegensatz zu „*improvement*“ meint „*enhancement*“ nicht eine Verbesserung eines (implizit) defizitären Zustands, sondern den Prozess der Qualitätsentwicklung (vgl. *Crozier/Curvale/Dearlove/Helle/Hénard 2008, S. 9*): „*enhancement is the process by which higher education providers systematically improve the quality of provision and the ways in which students' learning is supported*“, oder in der Terminologie der QAA „*taking deliberate steps at institutional level to improve the quality of learning opportunities*“ (vgl. u.a. QAA, 2011a). Die Philosophie von „*enhancement*“ ist dabei ganz klar, dass zu keiner Zeit ein Endpunkt von Qualitätsentwicklung erreicht ist – auch wenn die Lernmöglichkeiten, die eine Hochschuleinrichtung anbietet, anerkanntermaßen von hoher Qualität sind, kann diese immer noch weiter und innovativ entwickelt und vorangetrieben werden. Hochschuleinrichtungen müssen demzufolge Strukturen und Prozesse etabliert haben, die es ihnen erlauben, Defizite in der Qualität ihrer *learning opportunities* sowohl zu identifizieren als auch zu beheben. Zugleich sollen Hochschuleinrichtungen demonstrieren können, dass sie ein Bewußtsein für die Notwendigkeit beständiger Qualitätsentwicklung haben sowie mittels interner Strukturen und Prozesse Potentiale zur Qualitätsentwicklung erkennen und realisieren können. In diesem Sinne meint „*enhancement*“ vor allem auch einen die Hochschuleinrichtung insgesamt übergreifenden, strategischen Ansatz zur Qualitätsentwicklung. Idealerweise ist Qualitätsentwicklung immer auch Teil effektiver Qualitätssicherung.

Teil B des UKQCHE setzt die Referenzpunkte, anhand derer Bildungsanbieter die Qualität der von ihnen angebotenen Lernmöglichkeiten sichern und weiterentwickeln.⁵ Teil C des UKQCHE schließlich ist ein neuer Referenzpunkt hinsichtlich der Information, die Hochschuleinrichtungen über sich selbst und ihre Bildungsangebote geben. Dieser Teil des UKQCHE ist vor allem vor dem Hintergrund des politischen Diskurses um die informierte Selbstbestimmung Studierender zu sehen und zu verstehen; er formuliert eine Erwartung an den britischen Hochschulsektor hinsichtlich der Qualität jeglicher Information, die Hochschuleinrichtungen über sich und ihre Lehre veröffentlichen.

5 Die Begutachtungsverfahren der QAA im Überblick

Die QAA nimmt keine Programmakkreditierungen vor. Einrichtungen, die das Recht haben, akademische Abschlüsse und Grade zu vergeben (dies wird als *degree-awarding powers* bezeichnet), entwickeln, implementieren und – wenn sie von Partner-

⁵ Siehe <http://www.qaa.ac.uk/AssuringStandardsAndQuality/quality-code/Pages/Quality-Code-Part-B.aspx>, Zugriff 10.01.2012

organisationen/-einrichtungen gelehrt werden sollen – validieren Studienprogramme und -angebote in eigener Verantwortung. Die Begutachtungen der QAA prüfen, ob und wie britische Bildungseinrichtungen, die Studierende auf Universitätsniveau und für britische akademische Abschlüsse ausbilden, die akademischen Standards ihrer Bildungsprogramme und Qualifikationen wahren und sichern, und wie sie die Qualität ihres Lehr- und Lernangebotes kontinuierlich weiterentwickeln. Letztere bezeichnet die QAA als „*quality of learning opportunities*“. Die – wörtlich übersetzt – „Lernmöglichkeiten“ umgreifen alles, was eine Hochschuleinrichtung ihren Studierenden zur Verfügung stellt, um sie zur Erlangung eines akademischen Abschlusses zu befähigen. „*Learning opportunities*“ sind die akademische Lehre per se, Beiträge der Studierenden zu ihrem eigenen Lernen, sowie die akademische und soziale/persönliche Unterstützung, die Studierende während ihres Studiums erhalten und die es ihnen ermöglicht, Fortschritte in ihrem Studium zu erlangen. Ressourcen wie IT-Infrastruktur, Bibliotheken, Zulassungsregulierungen und gezielte Personalentwicklung für die akademische Lehre tragen ebenso zur Qualität der *learning opportunities* bei wie die Gestaltung eines Studienprogramms per se. Die Berichte jeder Begutachtung zeigen auf, wie Hochschuleinrichtungen ihrer jeweiligen Verantwortung für Standards und Qualität gerecht werden, identifizieren gute und beispielhafte Praxis, und unterbreiten Vorschläge, wie der jeweilige Bildungsanbieter die Qualität seiner Lehr- und Lernangebote verbessern sowie innovativ weiterentwickeln kann.

Die Begutachtungen durch die QAA sind peer review Verfahren. Kernziel jeder Begutachtung ist die Überprüfung, ob Hochschulprogramme, die zu akademischen Abschlüssen führen, auf angemessenem Niveau entwickelt und angeboten werden, ob Studierende in angemessener Weise zur Erlangung dieser Qualifikationen befähigt werden, und ob Hochschuleinrichtungen in angemessener Form ihr Recht ausüben, diese Abschlüsse zu verleihen. Die Erwartungshaltung jeder Begutachtung durch die QAA ist, dass eine Hochschuleinrichtung die Effizienz und Effektivität ihres Managements akademischer Standards und Qualität anhand der folgenden fünf Fragen kritisch reflektiert:

- was tun wir für das Management akademischer Standards und für die kontinuierliche Weiterentwicklung der Qualität unserer Lehr- und Lernangebote;
- warum tun wir das;
- wie tun wir das;
- warum ist das für uns der beste Weg;
- woher wissen wir, dass unser Management akademischer Standards und unsere Qualitätsentwicklung funktionieren;
- wie können wir beides noch verbessern.

Daraus folgt, dass Qualitätssicherung nicht nur im engen Sinne auf eine Sicherung und Verbesserung der Qualität akademischer Lehre abzielt. Jede Begutachtung durch die QAA hat die inneruniversitären Verwaltungs-, Regulierungs-, Monitoring- und Managementprozesse zum Gegenstand, die notwendige Voraussetzung und Umfeld für eine hohe Qualität akademischer Lehre sind. Beispielsweise sind Strukturen zur Erhebung und Evaluierung von studentischem Feedback zu Lehrveranstaltungen Bestandteil dieser im weiteren Sinne Managementprozesse. Im Zuge einer Begutachtung schaut die QAA jedoch nicht zu allererst auf die Ergebnisse solcher Umfragen, sondern darauf, dass Strukturen existieren, um solche Umfragen überhaupt durchzuführen, um die Ergebnisse zu sammeln und zu evaluieren, und vor allem auf Aktivitäten, die auf das studentische Feedback eingehen und zur Weiterentwicklung der Qualität des Lehrangebots genutzt werden.⁶ Dieses Beispiel zeigt, dass die Begutachtungen durch die QAA nicht zuletzt auch dazu dienen, Bildungsanbieter bei ihrer Planung und Strategieentwicklung zu unterstützen: Insbesondere der für jede Begutachtung einzureichende Selbstbericht, der um die oben zitierten fünf Fragen kreist, animiert Bildungsanbieter oft, schon vor einer Begutachtung reflektierend anzumerken, welche Aspekte von Qualitätsmanagement und -entwicklung strategisch überdacht werden können. Hierin spiegelt sich auch die Philosophie sämtlicher QAA Begutachtungsverfahren, dass Qualitätssicherung und Qualitätsentwicklung Hand in Hand gehen, und letztlich jede Begutachtung auch der Innovation dienen muss. Insbesondere die in Schottland angewandte *Enhancement-Led Institutional Review* (ELIR) folgt diesem Ansatz.

Die Begutachtungsverfahren der QAA berücksichtigen, welche Art von Hochschuleinrichtung überprüft wird und unter welche Regionaladministration diese fällt. Gegenwärtig kommen insgesamt neun Verfahren zur Anwendung.⁷

⁶ Der National Student Survey (NSS) ist das statistische Instrument, um studentisches Feedback auf nationaler Ebene zu sammeln, <http://www.thestudentsurvey.com/>. Die Ergebnisse werden bis September 2012 jährlich und per Hochschule auf Unistats publiziert, <http://unistats.direct.gov.uk/english/Index.do?t=201310020622294>; ab September 2012 wird der überwiegende Teil der Daten, die derzeit auf Unistats abrufbar sind, in den sogenannten Key Information Sheets (KIS) veröffentlicht. Für mehr Informationen zu dieser Entwicklung siehe vor allem Higher Education Funding Council for England (HEFCE), <http://www.hefce.ac.uk/learning/infohe/kis.htm> sowie Higher Education Statistics Agency (HESA), <http://www.hesa.ac.uk/content/view/2286/393/>. Zugriff auf alle Links am 10.01.2012

⁷ Überblick und Links zu Detailinformationen über die entsprechenden Begutachtungsverfahren sind zu finden unter <http://www.qaa.ac.uk/InstitutionReports/types-of-review/Pages/default.aspx>, Zugriff 10.01.2012.

5.1 England und Nordirland

- Die *Institutional review* wurde im September 2011 eingeführt und wird auf Hochschuleinrichtungen angewendet, die aus staatlichen Geldern finanziert werden;
- Das *Audit of collaborative provision* überprüft, wie Bildungsanbieter mit *degree-awarding powers* die akademischen Standards ihrer Abschlüsse wahren, wenn sie von anderen Bildungsanbietern offeriert werden;
- Die *Integrated quality and enhancement review* (IQER) wird auf Hochschulprogramme angewendet, die an further education colleges gelehrt werden; die Nachfolgemethode *Review of higher education in further education* (RHEFE) wird gegenwärtig erarbeitet und soll im akademischen Jahr 2012/2013 eingeführt werden.

In England finden im Kalenderjahr 2012 erstmals *Reviews for educational oversight* (REO) statt. Ihr müssen sich unabhängige, privat finanzierte (for-profit und non-profit) Colleges unterziehen, die Studierende aus dem nicht-EU Ausland rekrutieren. Die britische Grenzschutzbehörde (UK Border Agency, UKBA) hatte 2011 neue Richtlinien für die Zuwanderung ins Vereinigte Königreich erlassen. Danach können potentielle Studierende aus dem nicht-EU Ausland nur dann ein Visum für einen Aufenthalt während der gesamten Dauer eines Studienprogramms und eine Arbeitserlaubnis für diesen Zeitraum erhalten (sogenannter Tier 4), wenn sie nachweislich an einer Bildungseinrichtung mit *highly trusted sponsor status* registriert sind. Der überwiegende Teil der allgemein bekannten Hochschulen und Colleges hat diesen Status bereits inne, z.B. aufgrund des Körperschaftsstatus als öffentlich-rechtliche Universität und auch, weil sie regelmäßige Begutachtungen durch Einrichtungen wie die QAA durchlaufen. Es existieren jedoch um die ca. tausend Bildungseinrichtungen in privater Hand, die bis dato keinem solchen Verfahren unterzogen wurden. Um den Status als *highly trusted sponsor* zu erlangen, müssen sich solche Bildungseinrichtungen einer Begutachtung unter anderem durch die QAA unterziehen.⁸

⁸ Für Details zur Zuwanderungsregelung für Studierende aus dem nicht-EU Ausland siehe UK Border Agency (UKBA), <http://www.ukba.homeoffice.gov.uk/business-sponsors/education-providers/>. Die UKBA veröffentlicht ein ständig aktualisiertes Register von Einrichtungen, die als *sponsors* internationale Studierende anwerben und immatrikulieren können, UKBA, 2012, Register of Sponsors Licensed Under the Points-based System, <http://www.bia.homeoffice.gov.uk/sitecontent/documents/employersandsponsors/pointsbasedsystem/registerofsponsorseducation>. Das Handbuch für Bildungsanbieter, die als Sponsoren für internationale Studierende agieren (wollen), ist aufrufbar unter UKBA (2011), Guidance for Sponsors under Tier 4 of the Points-Based System, <http://www.ukba.homeoffice.gov.uk/sitecontent/documents/employersandsponsors/pbsguidance/guidancefrom31mar09/sponsor-guidance-t4-050911.pdf?view=Binary>. Alle Links wurden aufgerufen am 10.01.2012.

5.2 Wales

- Institutional review wird seit dem akademischen Jahr 2009/2010 auf aus staatlichen Geldern finanzierte Hochschulen angewendet und entspricht im wesentlichen der gleichnamigen Methode in England und Nordirland.
- Review of higher education (HE) in further education (FE) (Wales) wird auf Hochschulprogramme angewendet, die an further education colleges gelehrt werden.

5.3 Schottland

- Die Enhancement-led institutional review wird auf staatlich finanzierte Hochschuleinrichtungen angewendet.

In England und Wales führt die QAA außerdem die Relizenzierung und Begutachtung von *Access Validating Agencies* (AVA) durch. Bildungsprogramme unter dem *Access to Higher Education*⁹ Programm werden in einem breiten Fächerspektrum und vorwiegend von further education colleges angeboten. Sie bereiten Erwachsene, die nicht über die notwendigen formalen Qualifikationen und die Hochschulzugangsberechtigung verfügen, auf ein Hochschulstudium vor. Die QAA lizenziert die Organisationen, die im *Access to Higher Education* Programm einerseits Kurse entwickeln, validieren und kontinuierlich überwachen, und andererseits mit Bildungseinrichtungen zusammenarbeiten, die diese Kurse dann anbieten. Sie hat Kriterien für die Validierung von *Access* Kursen entwickelt; unter anderem muss eine Hochschule bestätigen, dass der zur Validierung vorgeschlagene Kurs tatsächlich und angemessen auf ein Hochschulstudium vorbereitet. Validierte Kurse werden von der QAA anerkannt und sind in der *Access courses database* zugänglich. Im Zuge der *AVA relicencing and review* prüft die QAA, ob und dass AVAs den Lizenzvereinbarungen und Lizenzkriterien entsprechen.

Der *General Osteopathic Council* (GOsC) ist die gesetzliche Regulierungsbehörde für Osteopathen und Osteopathieausbilder. Die Behörde setzt Standards und Qualitätsrichtlinien sowohl für die Ausbildung von Osteopathen an sich als auch für die Führung und das Management von Einrichtungen zur Ausbildung von Osteopathen. Bildungsangebote, die den Vorgaben des GOsC entsprechen, gelangen in den Status der *Recognised Qualification*. Nur mit dieser geschützten Qualifikation können sich Absolventen beim GOsC registrieren und als Osteopathen im Vereinigten Königreich arbeiten. Die QAA ist Vertragsnehmer des GOsC und führt in dessen Auftrag Überprüfungen durch, nach deren erfolgreichem Abschluss die Bildungsprogramme einer Ausbildungseinrichtung für Osteopathen den Status der *Recognised Qualification* erlangen, beibehalten oder erneuern.

⁹ Siehe <http://www.accesstohe.ac.uk/>, Zugriff 10.01.2012

Die QAA führt ausserdem das *Audit of overseas provision* durch.¹⁰ Britische Hochschul-einrichtungen bieten ihre Programme zunehmend im Ausland an, z. B. in Kooperation mit dortigen Bildungsanbietern. Auch für solche im Ausland angebotenen Hochschulprogramme liegt die Verantwortung für Qualität und Standards bei der britischen Hochschuleinrichtung. Die QAA begutachtet diese Kooperationsprogramme und prüft, ob und in welchem Maße britische Anbieter ihren Verantwortungen nachkommen. Die QAA nimmt keine Akkreditierungen oder Qualitätsprüfungen der Kooperationspartner im Ausland vor, hat aber einen Referenzrahmen entwickelt, der britische Anbieter bei der Identifikation potentieller Kooperationspartner sowie bei der Entwicklung und Durchführung von Kooperationen unterstützt.¹¹

6 Institutional Review England and Northern Ireland (IRENI)¹²

Im Folgenden wird auf die *Institutional Review England and Northern Ireland* (IRENI) näher eingegangen, die die QAA im Herbst 2011 als Nachfolgermethode des Institutional Audit eingeführt hat. IRENI wird auf higher education institutions (HEIs) im Sinne der Definition von Universities UK, dem Interessenverband des britischen Hochschul-sektors, angewendet, also auf *universities, university colleges, specialist higher education institutions and other higher education colleges*.¹³

IRENI trägt den politischen Entwicklungen im englischen Hochschulsektor der letzten zwei bis drei Jahre Rechnung. Dabei sind Studierende mehr und mehr in den Mittelpunkt des Geschehens gerückt, vor allem vor dem Hintergrund der gestiegenen Studiengebühren¹⁴. In England gibt es einen klaren Auftrag an den Hochschulsektor, die Interessen der Studierenden entschiedener zu berücksichtigen und ins tägliche Geschäft zu integrieren. Das Weissbuch der Regierung, im Herbst 2011 unter dem Titel „*Students at the heart of the system*“ veröffentlicht (*BIS 2011b*), läßt keinen Zweifel daran, dass im britischen Hochschulwesen ein Paradigmenwechsel stattfindet. Danach ist ein wesentliches Charakteristikum des gegenwärtigen britischen Hochschulsektors die

¹⁰ <http://www.qaa.ac.uk/InstitutionReports/types-of-review/overseas/Pages/default.aspx>, Zugriff 10.01.2012

¹¹ Vgl. den UK Quality Code for Higher Education, B10, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/quality-code-B10.aspx>, Zugriff 10.01.2012

¹² Siehe <http://www.qaa.ac.uk/InstitutionReports/types-of-review/Pages/IRENI.aspx> für weiterführende und detaillierte Informationen und Materialien, Zugriff 10.01.2012

¹³ <http://www.universitiesuk.ac.uk/UKHESector/Pages/OverviewSector.aspx>, Zugriff 10.01.2012

¹⁴ Maßgeblich dafür war der Sachverständigenbericht zur Finanzierung des englischen Hochschulwesens unter Federführung von Lord Browne im Jahr 2010. Ein Gesetzentwurf in Folge des Weissbuchs war für Frühjahr/Sommer 2012 erwartet worden. Die Regierung hat jedoch das Gesetzgebungsverfahren bis auf weiteres vertagt: Es wird befürchtet, dass die Umfragewerte für die konservativ-liberale Koalitionsregierung erheblich weiter sinken, wenn nach den begonnenen, umstrittenen Reformen des öffentlichen Gesundheitswesens (NHS) ein weiterer wesentlicher Sektor des öffentlichen Dienstes grundlegend – und nach Ansicht vieler Repräsentanten über das gesamte politische Spektrum hinweg zu sehr – nach den Vorstellungen gesellschaftlicher und politischer Eliten reformiert werden soll.

informierte Selbstbestimmung von Studierenden, deren Interessen Dreh- und Angelpunkt von Hochschulentwicklung sein sollen. Diese Philosophie spiegelt sich in der gegenwärtigen Dreijahresstrategie der QAA – als erstes strategisches Ziel hat sich die QAA gesetzt *to meet students' needs and to be valued by them (QAA 2011b)* – und ergo in den zur Anwendung kommenden Methoden zur Qualitätsprüfung und -sicherung. IRENI ist um sechs Kernprinzipien aufgebaut, die zugleich die wesentlichen Neuerungen der Methode gegenüber der Vorläufermethode (*Institutional Audit*) darstellen.¹⁵

Studierende im Zentrum der Qualitätssicherung

Da Qualitätssicherung letztlich dazu dienen soll, dass Studierende grösstmöglichen Gewinn aus ihrer Zeit an einer Hochschuleinrichtung erzielen, muss jede Methode der Qualitätssicherung und -prüfung auch Studierende einbeziehen. In IRENI holen die Gutachter nun mehr und detailliertere Rückmeldungen und Meinungen von grösseren Studierendengruppen ein. Nach der Begutachtung und dem Ergebnisbericht durch die QAA müssen Bildungsanbieter einen Aktionsplan in Kooperation mit ihren Studierenden entwickeln. Studierende sind Mitglieder der Gutachterteams. Außerdem benennt jede Bildungseinrichtung einen ihrer Studierenden, der in Kooperation mit dem Gutachterteams insbesondere die Gutachterbesuche koordiniert.

Flexibilität

Qualitätssicherung muss die Interessen sowohl des Hochschulsektors als auch der Öffentlichkeit berücksichtigen. Gutachterberichte sollen nicht nur Qualitätsmanager in Hochschulen bei der Wahrung und dem Management akademischer Standards sowie bei Massnahmen zur Qualitätsentwicklung unterstützen. Sie sollen auch auf die Interessen und Anliegen der Öffentlichkeit eingehen, die, angeregt durch den politischen Diskurs der vergangenen drei Jahre, zunehmend kritische und gesteigerte Ansprüche an Hochschulen und deren Dienste stellen. In IRENI beinhalten Qualitätsgutachten weiterhin die Beurteilungen (judgements) des Gutachterteams im Detail. Ein neuer Aspekt in IRENI ist die Behandlung eines spezifischen Themas, das jährlich sektorweit festgelegt wird, z.B. „*the first year experience*“ im Jahr 2011/2012. Die themenspezifischen Ergebnisse und Befunde der IRENI-Gutachten für 2011/2012 werden zusammenfassend publiziert, inklusive einer Analyse und Empfehlungen, wie der Sektor auf diese thematischen Befunde reagieren kann und soll. Um rascher auf Anliegen der Öffentlichkeit und des Hochschulsektors reagieren zu können, finden IRENI-Begutachtungen nicht mehr in einem fixen Zyklus statt, bei dem der gesamte Sektor alle sechs Jahre Qualitätsbegutachtungen unterzogen wird. Stattdessen durchläuft jede Einrichtung die Qualitätsbegutachtung in Intervallen von derzeit sechs Jahren. Der Zeitraum für das IRENI-Begutachtungsverfahren wurde verkürzt.

¹⁵ Für einen Vergleich der IRENI Methode mit der Vorgängermethode des Institutional Audit, siehe QAA (2011e); weiterführendes Material und Detailinformationen zu IRENI unter <http://www.qaa.ac.uk/InstitutionReports/types-of-review/Pages/onlinebriefing.aspx>, Zugriff 10.01.2012

Klarheit

Die Gutachterberichte müssen sowohl für Qualitätsmanager an Hochschulen als auch für die breite Öffentlichkeit informativ und hilfreich sein. Deshalb werden Beurteilungen deutlicher und allgemein verständlich formuliert. Zudem werden die Gutachterberichte insgesamt kürzer gehalten, und es wird ihnen eine an die breite Öffentlichkeit gerichtete Zusammenfassung vorangestellt.

Erkennbare Anerkennung des Stellenwerts guter Information über Hochschulbildung

Insbesondere zukünftige Studierende sollen gut informiert entscheiden können, an welcher Bildungseinrichtung sie was studieren möchten. Deshalb sind Art, Umfang und Qualität der Informationen, die Bildungseinrichtungen über sich sowie ihre Lehr- und Lernangebote geben und zugänglich machen, zunehmend von Bedeutung. IRENI unterzieht deshalb auch die Informationen, die Bildungseinrichtungen der breiteren Öffentlichkeit verfügbar machen, einer Prüfung und wird ab akademischen Jahr 2012/2013 ein Urteil über die Qualität dieser Informationen abgeben.

Vergewisserung der Öffentlichkeit, dass akademische Standards existieren und gewahrt sind

Die Wahrung akademischer Standards ist von grösstem Interesse im britischen sozio-ökonomischen Gesamtkontext. Um das öffentliche Interesse an und Vertrauen in gute und gesicherte akademische Standards zu wahren, geben IRENI-Gutachten klare Beurteilungen ab, ob bzw. dass die Lehr- und Lernangebote einer Bildungseinrichtung die nationalen Mindeststandards erreichen.

Minimale administrative Belastung

Eine Begutachtung durch die QAA soll keine zusätzliche administrative Belastung für Bildungseinrichtungen darstellen. Deshalb sollen alle Dokumentationen elektronisch eingereicht werden und Besprechungen zwischen Bildungseinrichtung, Gutachtern und QAA so oft als möglich per Telefon- und/oder Videokonferenz gehalten werden. Das zuvor übliche dreitägige Vorbereitungstreffen zwischen Gutachtertteam und Hochschuleinrichtung ist nicht länger Teil des Begutachtungsverfahrens.

Das Gutachtertteam beurteilt, ob und wie der Anbieter

- a) die für und durch den britischen Hochschulsektor vereinbarten akademischen Mindeststandards für Hochschulqualifikationen setzt und wahrt (Referenzpunkt ist Teil A des UKQCHE);
- b) Studierenden Lehr- und Lernangebote unterbreitet (im Sinne der „*learning opportunities*“, vgl. Kapitel 5), die es Studierenden ermöglichen, diese oben genannten Qualifikationen zu erlangen (Referenzpunkt ist Teil B des UKQCHE);

- c) die kontinuierliche Weiterentwicklung seines Lehr- und Lernangebot effektiv plant und managt (Referenzpunkt ist Teil B des UKQCHE);
- d) Informationen über sich selbst in angemessener Weise öffentlich zugänglich macht sowie die Qualität dieser Informationen sichert und managt (Referenzpunkt ist Teil C des UKQCHE; derzeit gibt ein Kommentar im Gutachterbericht Auskunft, ob und wie eine Einrichtung dem Referenzpunkt gerecht wird; ab dem akademischen Jahr 2012/2013 wird dies ebenfalls den Status einer Beurteilung erlangen).

Gutachter sprechen folgende Beurteilungen aus:

Für a):

- „meets UK expectations for threshold standards“ (entspricht den Erwartungen hinsichtlich der Wahrung und Sicherung akademischer Standards, die in den Referenzpunkten des UKQCHE gesetzt werden) oder
- „does not meet UK expectations for threshold standards“ (die Bildungseinrichtung wird den Erwartungen des UKQCHE nicht gerecht).

Für b) und c), ab akademischem Jahr 2012/2013 auch für d):

- „commended“ (keinerlei Beanstandungen und eigentlich beispielgebend hinsichtlich der Qualität und ihrer Weiterentwicklung);
- „meets UK expectations“ (keinerlei Beanstandungen und entspricht den Erwartungen, die in den Referenzpunkten des UKQCHE gesetzt werden);
- „requires improvement to meet UK expectations“ (Anstrengungen seitens der Bildungseinrichtung sind notwendig, um den Erwartungen der Referenzpunkte im UKQCHE entsprechen zu können);
- „does not meet UK expectations“ (die Bildungseinrichtung wird den Erwartungen des UKQCHE nicht gerecht).

Gutachtertteams sprechen je Urteil Empfehlungen aus, wie eventuelle Defizite im Qualitätsmanagement der Bildungseinrichtung beseitigt werden können, und formulieren auch einen Zeitrahmen, in dem Fortschritte und Verbesserungen erreicht werden müssen. Begutachtungen identifizieren darüber hinaus gute, beispielhafte und übertragbare Praxis, bestätigen gegebenenfalls Fortschritte in der Umsetzung von Empfehlungen aus früheren Begutachtungen, und regen weitere Innovationen und Verbesserungen an. Jede Hochschuleinrichtung in England und Nordirland durchläuft etwa alle sechs Jahre eine IRENI-Begutachtung.

Die QAA plant die Begutachtungen, wickelt alle administrativen Vorgänge ab (abgesehen von den internen Abläufen, die eine Hochschuleinrichtung in Vorbereitung und Organisation der Begutachtung selbst vornehmen muss) und setzt die Gutachtertteams zusammen (siehe dazu *QAA 2011c und QAA 2011d*). Dazu pflegt und erweitert die

QAA kontinuierlich einen Pool von Gutachtern, die über breite und vielschichtige Erfahrungen und Expertise im Hinblick auf den britischen Hochschulsektor verfügen. Eine Gutachtergruppe besteht in der Regel aus fünf Personen, davon ist eine ein Studierender. Ein QAA-Mitarbeiter bzw. eine Mitarbeiterin ist diesem Team als Koordinator zugeteilt und hat Verantwortung für den Ablauf der Begutachtung nach der IRENI-Methode. Gutachter werden durch ein dezidiertes QAA-Training auf die bevorstehenden Begutachtungen vorbereitet. Diese je nach Methode ein- bis zweitägigen Trainings führt die QAA für jede der zur Anwendung kommenden Begutachtungsmethoden statt, nicht nur für IRENI.

Die zu begutachtende Hochschuleinrichtung benennt eine Person, die als *institutional facilitator* während des gesamten Begutachtungsverfahrens Ansprechpartner für den Koordinator bzw. die Koordinatorin ist und wesentlich zur Kommunikation zwischen Bildungseinrichtung, QAA und Gutachterteam beiträgt. Desgleichen benennt die Hochschuleinrichtung einen *lead student representative*, der in wesentliche Kommunikationen zwischen Gutachterteam und Hochschuleinrichtung einbezogen ist, Treffen mit dem Gutachterteam wahrnimmt, vor allem die interne Kommunikation zwischen Studierendenschaft und Hochschuladministration wahrt und gestaltet sowie den schriftlichen Selbstbericht der Studierendenschaft koordiniert.

Von einer zu begutachtenden Hochschuleinrichtung wird erwartet, dass sie sich mit den von der QAA im Internet zur Verfügung gestellten Informationen und Materialien vertraut macht und selbständig alle notwendigen Prozesse in die Wege leitet, die seitens der Einrichtung für das Begutachtungsverfahren notwendig sind (für eine Übersicht über Zeitraum und Prozesse einer IRENI-Begutachtung siehe *QAA 2011f*). Das Gutachterteam stattet der Hochschuleinrichtung dann ca. 16 Wochen vor der eigentlichen Begutachtung einen Vorbereitungsbesuch ab. Während der eineinhalb Tage dieses Besuchs haben Einrichtung und Gutachter Gelegenheit, das eigentliche Begutachtungsverfahren zu besprechen, eventuelle Missverständnisse über Verfahren und Methode auszuräumen, sowie sich über Abläufe und Dokumentationen zu verständigen.

Hochschuleinrichtungen reichen vor jeder IRENI-Begutachtung einen reflektierenden Selbstbericht ein, in dem sie darlegen, wie sie Standards setzen und wahren sowie Qualität managen und weiterentwickeln. Idealerweise kontextualisiert dieser Selbstbericht die spezifischen Qualitätssicherungsprozesse und -strukturen einer Einrichtung und zeigt auf wie, auf welche Weise und mit welcher Motivation eine Einrichtung ihr Qualitätsmanagement entwickelt. Ein elektronisches Portfolio von Managementdokumenten dient der Untermauerung des Selbstberichts sowie zur Dokumentation von Prozessen und Strukturen, die die Gutachter dann untersuchen. Außerdem ist ein von der Studierendenschaft verfasster Bericht einzureichen, der darlegt, wie Studierende Lehre und Lernen an der jeweiligen Hochschuleinrichtung, Qualitätssicherung und -management

sowie die Repräsentation studentischer Interessen und Anliegen wahrnehmen. Selbstbericht und Bericht der Studierendenschaft sollen die einzigen Dokumente sein, die eine Hochschuleinrichtung in Vorbereitung einer Begutachtung anzufertigen hat; alle anderen Dokumentationen, z. B. Prüfungsprozesse, Gremienberichte oder Lehrevaluationen sollen nicht dezidiert erst in Vorbereitung einer Begutachtung angefertigt werden, sondern bereits existieren und Strukturen und Prozesse nachvollziehbar machen, die eine Hochschuleinrichtung zum Zweck der Qualitätssicherung bereits etabliert hat.

Zwei Wochen nach der eigentlichen Begutachtung informiert die QAA die Bildungseinrichtung sowie den Higher Education Funding Council for England, oder, im Falle der Begutachtung einer nordirischen Hochschule, das nordirische Department for Employment and Learning über die wesentlichen Ergebnisse der Begutachtung. Sechs Wochen nach der Begutachtung erhält die Hochschule eine Arbeitsfassung des Gutachterberichts sowie eine Zusammenfassung aller für den Bericht verwendeter Dokumentationen. An dieser Stelle hat die Bildungseinrichtung Gelegenheit, die faktische Korrektheit des Berichts zu prüfen und gegebenenfalls die Berichtigung von Fakten und Daten durch die QAA zu erwirken. Zwölf Wochen nach der Begutachtung veröffentlicht die QAA den Gutachterbericht auf ihren Internetseiten.¹⁶ Spätestens sechs Monate nach der Begutachtung durch die QAA veröffentlicht die Hochschuleinrichtung auf ihren Internetseiten einen Maßnahmenkatalog zur Umsetzung eventueller Empfehlungen durch die Gutachter. Nach drei Jahren überprüfen QAA und Hochschule die Fortschritte in der Umsetzung von Empfehlungen; innerhalb von sechs Jahren nach der ersten findet die nächste Begutachtung durch die QAA statt.

Die QAA evaluiert die Begutachtungen und holt von den Hochschulen Rückmeldungen zum relevanten Begutachtungsvorgang ein. Für IRENI, die im September 2011 eingeführt wurde, liegen zum jetzigen Zeitpunkt noch keine solchen Evaluierungen vor, da die Verfahren noch laufen.

Während der Einführung der IRENI-Methode hat das Department for Business, Innovation and Skills (BIS) eine Konsultation zur Weiterentwicklung des nationalen Rahmens für die Regulierung von Qualitätssicherung durchgeführt (*BIS 2011c*). Für die externe Qualitätssicherung relevant sind darin dargelegte Überlegungen, Häufigkeit und Intensität externer Qualitätssicherung an den *track record* einer Einrichtung zu koppeln: Ist die Einrichtung z. B. seit geraumer Zeit etabliert und haben frühere externe Begutachtungen keine Zweifel am Qualitätsmanagement der Einrichtung aufkommen lassen, würde sich diese Einrichtung für weniger häufige und weniger intensive externe Begutachtungen qualifizieren. Zum Zeitpunkt des Verfassens dieses Artikels sind die Ergebnisse der Konsultationen sowie Stellungnahmen und Implementierungspläne

¹⁶ Die Berichte sind zugänglich unter <http://www.qaa.ac.uk/InstitutionReports/Pages/default.aspx>, Zugriff 10.01.2012

seitens BIS nicht bekannt. Es kann jedoch angenommen werden, dass ein eventueller Paradigmenwechsel hin zu einem *risk-based approach* in der externen Qualitätssicherung Implikationen für die Begutachtungsmethoden der QAA haben wird.

Literatur

An Independent Review of Higher Education Funding and Student Finance (The Browne Review) (2010), <http://www.bis.gov.uk/assets/biscore/corporate/docs/s/10-1208-securing-sustainable-higher-education-browne-report.pdf>

Crozier, Fiona; Curvale, Bruno; Dearlove, Rachel; Helle, Emmi; Hénard, Fabrice (2007), Terminology of quality assurance: towards shared European values? ENQA Occasional papers 12, http://www.enqa.eu/files/terminology_v01.pdf.

Department for Business, Innovation and Skills (BIS) (2011a), The government's response to Lord Browne's review, <http://c561635.r35.cf2.rackcdn.com/11-1046-govt-response-to-browne-review.pdf>

Department for Business, Innovation and Skills (BIS) (2011b), Students at the heart of the system, White Paper, <http://c561635.r35.cf2.rackcdn.com/11-944-WP-students-at-heart.pdf>

Department for Business, Innovation and Skills (BIS) (2011c), A new, fit-for-purpose regulatory framework for higher education, Technical consultation, <http://c561635.r35.cf2.rackcdn.com/11-1114-new-regulatory-framework-higher-education-consultation.pdf>

BIS Portal zu aller relevanten Dokumentation hinsichtlich der geplanten Reformen des englischen Hochschulsektors, <http://discuss.bis.gov.uk/hereform/>

Further and Higher Education Act 1992, <http://www.legislation.gov.uk/ukpga/1992/13/contents>

QAA (2011a), UK Quality Code for Higher Education: General Introduction, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Quality%20Code%20General%20introduction%20Dec11.pdf>

QAA (2011b), Review for education oversight: Handbook, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/REO-handbook.aspx>

QAA (2011c), Institutional review of higher education institutions in England and Northern Ireland. A handbook for higher education providers, http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/IR_Handbook_March11.pdf

QAA (2011d), Institutional review of higher education institutions in England and Northern Ireland, Operational description, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/ireni-operational-d.pdf>

QAA (2011e), Summary of Institutional review in England and Northern Ireland, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/IRsummary.pdf>

QAA (2011f), UK Quality Code for Higher Education, General Introduction, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Quality%20Code%20General%20introduction%20Dec11.pdf>

QAA (2012), UK Quality Code for Higher Education, Part C: Information about higher education provision. Draft for consultation, <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/Quality-Code-Part-C-Information-about-higher-education-provision.pdf>

House of Commons Innovation, Universities, Science and Skills Committee (Select Committee) (2009), Students and Universities, <http://www.publications.parliament.uk/pa/cm200809/cmselect/cmdius/170/170i.pdf>

Welsh Assembly Government (2008), Review of Higher Education in Wales Phase 1: Student Finance Arrangements. Report from the Chair of the Task and Finish Group, <http://wales.gov.uk/docs/dcells/publications/4.0%20Policies%20and%20Strategies/5.0%20Further%20and%20Higher%20Education/Further%20and%20Higher%20Education/Review%20of%20Higher%20Education%20in%20Wales%20Phase%201%20Student%20Finance%20Arrangements/081201reviewhigher.pdf>

Welsh Assembly Government (2009), For our future, The 21st Century Higher Education Strategy and Plan for Wales, <http://wales.gov.uk/docs/dcells/publications/091214hestrategy.pdf>

Welsh Assembly Government (2009), Review of Higher Education in Wales Phase 2: <http://wales.gov.uk/docs/dcells/publications/090622hephase2en.pdf>

Auf alle Internetlinks wurde zuletzt am 10.01.2012 zugegriffen.

Anschrift der Verfasserin:

Dr. Anett Löscher
Development Officer (Standards, Quality and Enhancement)
Quality Assurance Agency for Higher Education
Southgate House
Southgate Street
UK-Gloucester GL1 1UB
E-Mail: a.loescher@qaa.ac.uk

Dr. Anett Löscher arbeitet im Standards, Quality and Enhancement Team der Quality Assurance Agency am UK Quality Code for Higher Education (Kapitel B7, External Examining; Part C: Information about higher education provision; B10: Management of collaborative arrangements). Sie ist verantwortlich für den *employer-based training accreditation service*, der Arbeitgebern hilft, das akademische Potential ihres In-House Trainings zu erkennen, mit Kreditpunkten zu versehen, und durch Universitäten akkreditieren zu lassen (vgl. <http://ebta.qaa.ac.uk/>).

Wissenschaft aus Kalkül? Wissenschaftsinterne Antizipation von Wettbewerbsmechanismen im Rahmen von Evaluationsverfahren

Silke Gülker

Inwiefern können extern initiierte Evaluationsverfahren die etablierten Koordinationsmechanismen in der Wissenschaft verändern? Dieser Frage widmet sich der Beitrag konzeptionell und empirisch. Als traditionell in der Wissenschaft akzeptiert gilt eine Kombination aus Wettbewerb (Konkurrenz um die besten Ideen) und Gemeinschaft (kollegiale Kooperation), während Entscheidungen qua Hierarchie eine untergeordnete Rolle spielen. Anhand einer Analyse der Evaluationsverfahren von Instituten der Leibniz-Gemeinschaft in Deutschland und der Research Assessment Exercises in Großbritannien wird gezeigt, dass die akzeptierten Mechanismen durchaus hinterfragt, aber nicht grundlegend in Frage gestellt werden. Die Verfahren provozieren zwar strategisches Verhalten, die Normen der Wissenschaftsgemeinschaft behalten aber zugleich ihre maßgebliche Bedeutung. Im Rahmen der Vorbereitung auf Evaluationsverfahren entsteht die Notwendigkeit, neue Entscheidungen zu treffen, die hierarchisch gelöst werden.

1 Einleitung

Wenn auf einer wissenschaftlichen Tagung das Stichwort Evaluation aufgebracht wird, dann werden die Debatten schnell hitzig. In der wissenschaftlichen Fachgemeinschaft scheint weitgehender Konsens darüber zu bestehen, dass mit Evaluationen etwas in die Wissenschaft eingeführt wird, was ihr fremd und nicht angemessen ist und ihr damit letztlich Schaden zufügt. Als Befürchtung steht dann schnell im Raum, dass Wissenschaft zunehmend an der Logik von Markt und Wettbewerb orientiert werde – wiederum als wissenschaftsfremd und unangemessen bewertet.

Ein Zusammenhang zwischen Evaluation und Wettbewerb liegt dann nahe, wenn Evaluationen als wissenschaftspolitisches Instrument eingesetzt werden und im Ergebnis Allokationsentscheidungen bei knappen Ressourcen nach sich ziehen. Für den Evaluierten hängt dann von der Evaluation viel ab und es gilt, sich im Wettbewerb durchzusetzen. Gleichzeitig bleibt die Frage, was „sich durchsetzen“ in diesen Verfahren, die in aller Regel auf einem Peer Review Prozess aufbauen, genau bedeutet. Inwiefern verlangen solche Evaluationen von Wissenschaftlerinnen und Wissenschaft-

lern etwas anderes als es die zur Sicherung wissenschaftlicher Qualität akzeptierten Verfahren der Fachgemeinschaft tun?

Der vorliegende Beitrag thematisiert diese Fragen in der Begrifflichkeit von Koordinationsmechanismen. Ziel ist, das Verhältnis von extern initiiertem Evaluation und wissenschaftsinternen Koordinationsmechanismen näher zu bestimmen. Traditionell wird für die Wissenschaft eine spannungsreiche Kombination aus Wettbewerb (Konkurrenz um die besten Ideen) und Gemeinschaft (Kooperation unter Fachkollegen) angenommen, während Hierarchie als Koordinationsmechanismus als weniger bedeutsam eingeschätzt wird. Inwiefern können Evaluationsverfahren diese etablierten Mechanismen in Frage stellen, inwiefern also werden externe Wettbewerbsmechanismen wissenschaftsintern antizipiert? Die Verfahrensregeln von Evaluation werden dabei als wesentliche Bedingung angenommen, das heißt, es wird davon ausgegangen, dass mit unterschiedlichen Verfahrensregeln auch unterschiedliche Wirkungen auf die Koordinationsmechanismen in der Wissenschaft verbunden sind.

Die Frage steht im Zusammenhang mit grundlegenden Auseinandersetzungen der Wissenschaftsforschung. Auf dem Spiel steht nichts weniger als die Autonomie der Wissenschaft. Wenn Evaluationen in der Lage sein sollten, interne Koordinationsmechanismen der Fachgemeinschaft grundlegend in Frage zu stellen, wäre damit auch eine Nejustierung im Verhältnis von Politik und Wissenschaft verbunden.

Einen Beitrag zur Beantwortung der Fragestellung soll eine empirische Analyse zweier sehr unterschiedlich organisierter Evaluationsverfahren leisten: der Evaluation von Instituten der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL) in Deutschland und der Research Assessment Exercises (RAE) in Großbritannien. Kern der Evaluationsverfahren in der Leibniz-Gemeinschaft ist eine Begehung der zu bewertenden Institute durch eine Gutachterkommission. Die RAE dagegen basiert im Kern auf der Begutachtung von Zeitschriftenartikeln aller so genannter forschungsaktiven Wissenschaftlerinnen und Wissenschaftler einer Forschungseinheit. Beide Verfahren bilden am Ende die Grundlage für Entscheidungen zur weiteren Förderung, sind also politisch initiierte Bewertungsverfahren im Wettbewerb um knappe Ressourcen¹.

Im Folgenden werden zunächst Datenbasis und methodisches Vorgehen dargelegt. In Abschnitt 3 wird nachvollzogen, welche spezifische Kombination von Koordinationsmechanismen für die Wissenschaft traditionell angenommen wird, um sodann potenzielle Einflusspotenziale durch Evaluationen spezifizieren zu können. Abschnitt 4 und

¹ Diese Verknüpfung von Evaluation und Mittelallokation ist in der Wissenschaftspolitik in den letzten Jahren international bedeutsamer geworden (*Power 2008, Whitley 2010*), und auf diese spezifische Form der Evaluation fokussiert hier die Analyse. Gleichzeitig können Evaluationen in der Wissenschaftspolitik wie auch in anderen Politikfeldern auch mit dem primären Ziel der Qualitätsentwicklung eingesetzt werden (*Kromrey 2003*).

5 präsentieren jeweils anhand von Verfahrensregeln und Interviewaussagen die empirischen Ergebnisse. Abschnitt 6 fasst die Aussagen in einem Fazit zusammen.

2 Datenbasis und methodisches Vorgehen

Die Untersuchung ist Teil des am Wissenschaftszentrum Berlin für Sozialforschung durchgeführten Projekts „Urteilsbildung im Peer Review. Internationale Fallstudien“. Das Projekt befasst sich mit der übergeordneten Fragestellung, wie Gutachterbewertungen in unterschiedlichen Verfahrenskontexten zustande kommen und welche Implikationen damit für die Bewerteten verbunden sind. Mit den Verfahren der WGL in Deutschland und nach dem Standard Evaluation Protocol (SEP) in den Niederlanden auf der einen Seite und der RAE in Großbritannien auf der anderen Seite wurden Prozesse in Verfahrenskontexten untersucht, die maximal kontrastieren. Die je Kontext ausgewählten Evaluationsfälle (ein Fall ist ein institutsbezogenes Verfahren) sind sowohl in Bezug auf die Fächergruppen als auch in Bezug auf die Institutgröße heterogen.

Weil Gutachterverfahren prinzipiell hinter verschlossenen Türen stattfinden (vgl. z. B. auch *Lamont 2009*), galt es eine Materialbasis zu schaffen, die eine weitgehende Rekonstruktion einzelner Begutachtungsprozesse ermöglicht. Nach explorativen Experteninterviews zur Entwicklung des Erhebungskonzeptes wurden in der Regel² zeitnah vor der Begutachtung Interviews mit Verantwortlichen der Institute geführt, nach Abschluss der Begutachtung dann mit Gutachterinnen und Gutachtern und nach Veröffentlichung der Ergebnisse noch einmal mit den Verantwortlichen der Institute. Die Tabelle illustriert, dass für jeweils drei Fälle pro Verfahrenskontext eine umfassende Erhebung in diesem Sinne möglich war. Die im Vergleich höhere Anzahl an Interviews im deutschen Kontext geht auf forschungspraktische Gründe zurück – die Erhebung hat in Deutschland begonnen und im Verlauf wurde eine Reduktion erforderlich.

² Für einzelne Fälle der WGL musste der Vorbereitungsprozess in einem Interview rekonstruiert werden, das erst nach der Begehung stattfinden konnte.

Tabelle: Übersicht der im Gesamtprojekt geführten Interviews

Codename Institut	Fachgruppe	Interviews mit/im				gesamt
		Institut vorher	Gutachter /innen	Institut nachher	Koordina- tion	
Verfahren der WGL in Deutschland						
Alpha-D	Geisteswissenschaften	2	2	2	1	7
Beta-D	Natur- und Sozialwissenschaften	7	5	7	1	20
Gamma-D	Natur- und Sozialwissenschaften	7	5	2	1	15
Delta-D	Natur- und Ingenieurwissenschaften	2	6	2	1	11
Epsilon-D	Naturwissenschaften	1	4	1	1	7
Verfahren nach dem SEP in den Niederlanden						
Alpha-NL	Geisteswissenschaften	2	5	2	1	10
Beta-NL	Sozialwissenschaften	2	2	2	1	7
Gamma-NL	Naturwissenschaften	2	4	1	1	8
Verfahren nach der RAE in Großbritannien						
Alpha-UK	Geisteswissenschaften	2		1		3
Beta-UK	Natur- und Sozialwissenschaften	3	3	1	1	8
Gamma-UK	Natur- und Ingenieurwissenschaften	3	3	1	1	8
Epsilon-UK	Natur- und Sozialwissenschaften	2	3	1	1	7
Fallbezogene Interviews gesamt		35	42	23	11	111
Explorative Expertengespräche						13
Interviews gesamt						124

Das Interviewmaterial dient im Projekt als Basis für vielfältige spezifische Auswertungen und wird je nach Fragestellung durch Dokumentenanalysen ergänzt³. Aus Gründen der Komplexitätsreduktion konzentriert sich der folgende Beitrag auf Fälle in Deutschland und in Großbritannien. Die Fragestellung bezieht sich auf die Antizipation von Evaluationsverfahren innerhalb der beteiligten Forschungseinrichtungen, die insgesamt 47 Interviews in den Instituten vor und nach der Begehung bilden entsprechend die Materialbasis. Interviewsegmente mit Bezug auf Koordinationsmechanismen wurden sequenziell analysiert (*Soeffner 1989, Wernet 2006*) und damit Deutungsmuster im Sinne von Interpretationsmustern (*Lüders/Meuser 1997*) rekonstruiert.

³ So befassen sich Beiträge konzentriert mit der Rolle und dem Rollenverständnis der Gutachter (*Torka 2011a*), mit genderspezifischen Fragen (*Matthies/Matthäus 2010*), mit dem Handlungsmodus in Begutachtungsprozessen (*Gülker/Simon/Torka 2012*) oder mit den institutionellen Folgen der Bewertungen (*Gülker/Simon 2011, Torka 2011b*).

3 Inwiefern können Evaluationen die Koordinationsmechanismen der Wissenschaft verändern?

Von Koordinationsmechanismen in der Zuspitzung auf die Mechanismen Markt, Gemeinschaft und Hierarchie zu sprechen, ist stets mit Vereinfachungen verbunden, die Reduktion auf solche in Reinform niemals existenten Koordinationstypen lässt vielfältige Dimensionen zur Erklärung sozialer Prozesse außer Acht. Gleichzeitig wird insbesondere in der Kritik an jüngeren Strategien der Wissenschaftspolitik immer wieder mit Begriffen wie Markt und Wettbewerb argumentiert; damit verbunden werden explizit oder implizit Annahmen eben in Bezug auf die Koordinationsmechanismen der Wissenschaft. So ist beispielsweise die Rede vom „Niedergang der Kollegialität“ (Martin/Whitley 2010), von „Forschung und Lehre aus Kalkül“ (Kieser 2010), von „Verdrängung intrinsischer Arbeitsanreize“ und „Manipulation der Kennziffern“ (Frey 2008). Im Sinne einer pointierten Auseinandersetzung mit diesen Thesen soll im folgenden vereinfacht davon ausgegangen werden, dass sich Koordination in der Wissenschaft wesentlich anhand dieser Typisierung von Mechanismen erklären lässt. Und diverse Arbeiten der Wissenschaftssoziologie lassen es zu, in der Wissenschaft eine Kombination aus Wettbewerb und Gemeinschaft als dominierende Mechanismen anzunehmen, während Organisation und Hierarchie generell eine geringere Bedeutung beigemessen wird.

Der besondere Mechanismus des *Wettbewerbs* ist, dass mehrere Beteiligte miteinander konkurrieren, dabei aber im Prinzip das gleiche Ziel verfolgen. Konkurriert wird beispielsweise um knappe Ressourcen, um den besten Platz im Sport oder um die meisten Wählerstimmen in der Politik (Benz 2007). Weil der eigene Erfolg immer in Relation zum Erfolg der anderen steht, sind die Konkurrenten zur Bestleistung angereizt. Die Koordination erfolgt über gegenseitige Beobachtung und Anpassung im Sinne von Anstrengungen zur Überbietung der Konkurrenz. Im positiven Modell wird im Wettbewerb so ein „optimales“ Ergebnis „entdeckt“ (Hayek 1969). Grenzen der Optimierung durch Wettbewerb liegen dort, wo die volle Konzentration auf das Wettbewerbsziel auf Kosten anderer Zielsetzungen geht. Regeln müssen verhindern, dass Konkurrenten sich untereinander Schaden antun, Ziele und Maßstäbe des Wettbewerbs müssen gut definiert sein. Setzen die Regeln der individuellen Nutzenmaximierung keine Grenzen, dann werden auch List und Täuschung angewandt, um den eigenen Vorteil zu realisieren.

Der Koordinationsmechanismus *Gemeinschaft* kontrastiert in gewissem Sinne mit dem des Wettbewerbs. Nicht das Ziel, anders im Sinne von besser zu sein als die anderen, treibt hier die Entscheidungs- und Handlungsprozesse, sondern die positive Identifikation mit der Gemeinsamkeit mit anderen. Der Gemeinschaftsbegriff hat eine lange Tradition und wurde im Laufe des vergangenen Jahrhunderts auf zum Teil sehr

unterschiedliche soziale Gebilde angewandt (*Gläser 2007*). Der gemeinsame Ausgangspunkt der unterschiedlichen Verwendungen ist die Vorstellung, dass Mitglieder ihr Handeln auf Grundlage geteilter Werte koordinieren. In unterschiedlicher Betonung zeichnen sich Gemeinschaften durch vorgängig gegebene Zugehörigkeiten und Bindungen aus, die nicht – im Unterschied zur Gesellschaft – ausgehandelt oder rational entschieden werden. Durch Beobachtung und Anpassung entwickeln sich tief verwurzelte implizite oder explizite Regeln, die für die Mitglieder handlungsleitend werden.

Zur Erklärung wissenschaftlicher Dynamik spielt nun ein spezifisches, spannungsreiches Verhältnis von Wettbewerb und Gemeinschaft traditionell eine wichtige Rolle. Wettbewerb um die beste Idee, das beste Argument wird schon von Popper als wesentlicher Antrieb für wissenschaftliche Leistungen eindrücklich beschrieben – ebenso wie die stete Gewissheit, dass der eigene Beitrag zwangsläufig früher oder später von den Kollegen überboten werden wird (*Popper 1959*). Im Konzept von Bourdieu (*1975*) spielt Konkurrenz um Reputation die zentrale Rolle bei der wissenschaftlichen Wissensproduktion. Er benutzt ebenso wie schon Hagstrom (*1974*) ökonomische Begriffe wie Tausch und Investition zur Erklärung wissenschaftlicher Dynamik. Latour und Woolgar (*1986*) beschreiben dann auf der Grundlage ihrer Laborstudien ein ökonomisches Modell, das zusätzlich erklären soll, warum sich Wissenschaftlerinnen und Wissenschaftler stets auf die Arbeiten der anderen beziehen müssen.

Wettbewerb an sich ist also als integraler Bestandteil der Wissenschaft akzeptiert. Dieser Wettbewerb folgt allerdings spezifischen Regeln. Als etablierte (Wettbewerbs-) Regeln in der Wissenschaft kann man solche annehmen, die mit den Normen der Wissenschaftsgemeinschaft übereinstimmen. Dafür können die von Merton (*1973*) aufgestellten Normen nach wie vor als handlungsleitend angenommen werden. Insbesondere das Prinzip des „organisierten Skeptizismus“ wird auf allen Ebenen wissenschaftlicher Arbeit praktiziert. Als Koordinationsmechanismus lässt sich dieses Prinzip als „kollegiales Fachurteil“ beschreiben. Die Arbeit der anderen Gemeinschaftsmitglieder wird zum einen bewertet vor dem Hintergrund der eigenen Kenntnis und Deutung des Fachgebietes und zum anderen auch stets im Wissen darum, dass bei nächster Gelegenheit die eigene Arbeit dem Urteil der anderen ausgesetzt sein wird. Diese gegenseitige Abhängigkeit gibt einen starken Anreiz, Urteile gut (im Sinne von fachkompetent) inhaltlich zu begründen, fachfremde Eigeninteressen nicht (jedenfalls nicht erkennbar) in die Bewertung einfließen zu lassen und zu scharfe Formulierungen zu unterlassen (*Neidhardt 1988: 114ff*). Kollegiale Fachurteile werden dann akzeptiert, wenn sie den Normen der Wissenschaftsgemeinschaft entsprechen.

Koordination qua *Hierarchie* wird traditionell in der Wissenschaft als weniger bedeutsam angenommen. Hierarchie meint grundlegend zunächst Verfahren der Über- und

Unterordnung und ist ein Teilaspekt des Mechanismus *Organisation* (vgl. *Döhler 2007*). Organisation betont generell die Entscheidungsbedingtheit von Sachverhalten (*Wiesenthal 2000*) und setzt im wesentlichen auf klare Zuständigkeitsregeln. Im Idealtyp einer hierarchischen Bürokratie von Weber (*1980 [1922]*) werden bereits die wesentlichen Aspekte hervorgehoben, an denen sich auch heutige Untersuchungen orientieren, Organisationen werden anhand ihres Ausmaßes an Arbeitsteilung, Amtshierarchie, technischer Regeln und Normen sowie Aktenmäßigkeit klassifiziert.

Für die Wissenschaft wird weithin ein Organisationstypus angenommen, in dem Hierarchie als Koordinationsmechanismus nicht stark ausgeprägt ist. Damit ist nicht gemeint, dass es in der Wissenschaft keine Statusunterschiede geben würde, diese Unterschiede werden aber nur dann akzeptiert, wenn sie Ergebnis oben genannter Prozesse von Wettbewerb und Gemeinschaft sind. Hierarchie als Entscheidungsmechanismus wird als ungeeignet angenommen, weil es sich bei wissenschaftlichen Organisationen in der einen oder anderen Weise um „unvollständige“ (*Brunsson/Sahlin-Andersson 2000*) Organisationen handelt. Bei Mintzberg (*1979*) etwa gelten wissenschaftliche Einrichtungen als Idealtyp einer „Professional Bureaucracy“, geprägt von großer Abhängigkeit von Fachleuten, deren Standards auch in der Fachgemeinschaft und nicht in der Organisation gesetzt werden. Weil Ziele in der Wissenschaft nicht konsistent und nicht operationalisierbar sind, Wissen und Technologien beschränkt sowie die Teilnehmerschaften wechselnd sind, geben Cohen et al. (*1972*) der Universität den Titel einer „organisierten Anarchie“. Diese Zuschreibungen argumentieren mit der Logik wissenschaftlicher Arbeit als hierarchisch nicht organisierbar und prägen die Strukturen in der Wissenschaft nachhaltig.

Inhaltliche Konkurrenz und normbasiertes gemeinschaftliches Handeln sind also die prägenden Mechanismen der Fachgemeinschaften, während Entscheidungen qua Hierarchie eine untergeordnete Rolle spielen. Die folgende empirische Analyse prüft, inwiefern Evaluationsverfahren in der Lage sein können, diese Kombination von Mechanismen in Frage zu stellen. Untersucht werden dafür:

1. *Die Ausgestaltung der Wettbewerbsregeln*: Wenn der durch Evaluationen organisierte Wettbewerb mit den Normen der Fachgemeinschaft nicht übereinstimmen würde, wäre damit der übliche Koordinationsmechanismus irritiert. Dies wäre der Fall, wenn die Regeln Signale belohnen würden, die von wissenschaftlichen Inhalten entkoppelt wären.
2. *Das Verhalten der Beteiligten*: Nicht den üblichen Mechanismen entsprechen würde ein strategisches Verhalten, das ansonsten von der wissenschaftlichen Fachgemeinschaft sanktioniert werden würde. Dazu gehörte die Anwendung von List und Täuschung, um die inhaltliche Leistung besser dastehen zu lassen oder schlechte

Leistung zu vertuschen. Dazu gehörte weiterhin, wenn Entscheidungen qua Hierarchie an Bedeutung und Akzeptanz gewinnen würden.

4 Die Evaluation der Leibniz-Gemeinschaft

4.1 Verfahrensregeln

Die Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz wurde 1997 als Nachfolgeorganisation der „Wissenschaftsgemeinschaft Blaue Liste“ gegründet. Ihr gehören derzeit 87 Forschungsinstitute an, die von Bund und Ländern gemeinsam finanziert werden. Einrichtungen der Blauen Liste werden bereits seit 1979 im Auftrag von Bund und Ländern regelmäßig evaluiert, verantwortlich ist der Senatsausschuss Evaluierung (SAE), eingesetzt vom Senat der Leibniz-Gemeinschaft. Jedes Mitgliedsinstitut wird alle sieben Jahre evaluiert. Für jede Institutsevaluation wird ein eigenes Begutachtungsgremium eingesetzt, zusätzlich sind Delegierte von Bund und Länder als Gäste an den Verfahren beteiligt. Kern des Bewertungsprozesses ist ein Vor-Ort-Besuch von eineinhalb Tagen, die so genannte „Begehung“, vorbereitet durch eine standardisierte schriftliche Selbstdarstellung des Instituts. Über die Begehung wird ein Bericht verfasst. Zu diesem Bericht kann die bewertete Einrichtung eine Stellungnahme abgeben, bevor der Senat der Leibniz-Gemeinschaft seine Stellungnahme mit wissenschaftspolitischen Empfehlungen formuliert. In letzter Instanz dient dieser Bericht der Gemeinsamen Wissenschaftskonferenz zur Überprüfung der gemeinsamen Förderung von Bund und Ländern.

Inwiefern legt der Verfahrensrahmen es nahe, dass etablierte Koordinationsmechanismen der Wissenschaft in Frage gestellt werden? Zunächst ist auffällig, dass Initiative und Regelwerk wesentlich von Mitgliedern der Fachgemeinschaft getragen werden. Zwar handelt es sich auch hier um einen wissenschaftspolitisch gestalteten Prozess – am Ende steht eine Entscheidung über die Förderfähigkeit des Instituts. Aber nicht nur die Peers sind an zentraler Stelle positioniert und mit erheblichem Gestaltungsspielraum ausgestattet, auch der SAE als verantwortliches Gremium ist überwiegend mit Wissenschaftlerinnen und Wissenschaftlern besetzt. Das Vorgehen ist folglich nah an den Routinen der Fachgemeinschaft orientiert: Am Ende zählt das in der Interaktion der Gutachterinnen und Gutachter untereinander abgewogene Urteil, ausgedrückt in einem ausführlichen Bericht – ein Vorgang, der für Wissenschaftlerinnen und Wissenschaftler selbstverständlich ist.

Nicht alltäglich allerdings ist die Ebene und damit der Gegenstand der Bewertung: In dem Verfahren werden Institute als Ganze auf den Prüfstand gestellt. Von den insgesamt acht Überschriften der Kriterienliste bezieht sich eine auf die Arbeitsergebnisse, also die Ebene, die üblicherweise unter Kollegen zur Diskussion gestellt wird. Die anderen Überschriften heißen *„Gesamtkonzept und Arbeitsschwerpunkte“*,

„Kooperation und wissenschaftliches Umfeld“, „Nachwuchsförderung“, „Struktur und Management der Einrichtung“, „Mittelausstattung und -verwendung“, „Personal“ und „Umsetzung der Empfehlungen der letzten Evaluierung“ (Leibniz-Gemeinschaft 2007). Hier konkurrieren also Institute auch in Bezug auf ihre Managementkompetenzen und nicht allein in Bezug auf wissenschaftliche Inhalte miteinander.

4.2 Antizipation in den Instituten

Als strategisches im Sinne von wettbewerblichem Verhalten läge nun auf Seiten der Institute nahe, alles zu tun, um das gesamte Institut gut dastehen zu lassen. Offene Fragen und Probleme gilt es möglichst zu vertuschen – anders als dies jedenfalls bei einem inhaltlichen Feedback unter Kollegen der Fall sein muss. Die Darstellungen der Institutsvertreterinnen und -vertreter machen in der Tat deutlich, dass die Evaluation als ein erheblicher Druck im Sinne einer Prüfungssituation wahrgenommen wird. Der Vergleich mit anderen Instituten spielt dabei eine untergeordnete Rolle. Am Ende des Verfahrens steht kein Ranking ähnlicher Institute, sondern es geht vielmehr darum, das eigene als besonders und so spezifisch darzustellen, dass es auch weiterhin als förderungswürdig akzeptiert wird. Und in diesem Zusammenhang werden die Gutachter, die an anderer Stelle vor allem fachlich geschätzte Kollegen sind, auch zu einem Gegenüber, denen es aus strategischen Gründen gilt, einen guten Aufenthalt zu gestalten:

„Die Leute waren also würde ich mal sagen alle gut gelaunt. Das ist glaube ich auch schon mal wichtig [lacht] [...]. Wir haben uns sehr bemüht, was jetzt auch so Transportdienst und so was angeht also hier, hatten einen Shuttle-Service organisiert und wir haben hier einen Kollegen, der meiner Ansicht nach wunderbar die Rundum-Sorglos-Pakete vorbereitet hat für die Leute, dass sie halt wirklich an nichts denken müssen und auch genau wissen, wo sie wann stehen müssen und solche Geschichten. (Bock, Beta-D, 644)⁴

Die Darstellung unterstreicht, dass die Begehung als wichtiges Ereignis angenommen wird, bei dem nichts schief gehen darf. Ähnliche Aktivitäten werden in allen Instituten beschrieben, sei es in Bezug auf die Raumgestaltung, Sprecherwahl und -wechsel oder Posterpräsentationen. Abgesehen von diesen durchaus als Eventmanagement zu charakterisierenden Anteilen sehen sich die Institute durch die Evaluation aber auch gefordert, sich mit ihrer Organisation systematisch auseinander zu setzen – allein die gute Gestaltung der Fassade wird nicht als adäquat angenommen. Das Evaluationsverfahren wird vielmehr als eine Art erzwungene Organisationsberatung antizipiert. In

⁴ Alle Namen der Interviewpartnerinnen und -partner wurden in einem Anonymisierungsverfahren geändert. Die Referenzen bezeichnen 1. Codenamen der Interviewten, 2. Codename des Instituts (vgl. Tabelle 1), 3. Startzeile der Interviewtranskription.

allen befragten Instituten setzt der Vorbereitungsprozess auf die Evaluation mindestens ein Jahr im Voraus ein und beinhaltet eine Überprüfung der eigenen Struktur. Beispielhaft sei hier die Darstellung eines Interviewpartners, Herrn Müller aus Beta-D, nachvollzogen.

Herr Müller ist stellvertretender Direktor des Instituts, und er bewertet den Evaluationsprozess insgesamt vor allem deshalb positiv, weil seitens der Leibniz-Gemeinschaft nicht nur nach vergangenen Leistungen gefragt wird, sondern auch die Zukunftsperspektiven bewertet werden. Um diese Perspektive zu entwickeln, wurde auf Anraten des wissenschaftlichen Beirats ein zweitägiger Workshop organisiert, an dem *„[...] ganz ausgewählte 25 Personen im Institut einmal am Wochenende diskutiert haben und daraus hat sich dann die Idee kristallisiert, die Herr Uckermann [Leiter des Instituts, Anm. S.G.] dann vollständig in Text gegossen hat.“* (Müller, Beta-D, 83) Dieser Workshop *„[...] war dann definitiv ne Vorbereitung auf die Zukunft und natürlich dann auf die Evaluierung.“* (Müller, Beta-D, 91)

Die Evaluation ist also Anlass für interne Selbstverständigungsprozesse, denen über die Evaluation hinaus Bedeutung gegeben wird. Herausgefordert werden solche Selbstverständigungsprozesse durch den Zwang zur Verschriftlichung. So beschreibt Herr Müller, dass bei der Formulierung des Institutsprofils eine Diskussion über das Verhältnis von Politikberatung und Grundlagenforschung entstanden sei. Über dieses Grundverständnis hat dann ein schriftlicher Austausch stattgefunden. *„Das war dann spannend, das hat auch ich denke ne Menge an Selbsterkenntnis für das Institut gebracht“* (Müller, Beta-D, 130). Unumstritten war die dann verschriftlichte Grundidee einer Verbindung von Politikberatung und Grundlagenforschung nicht:

„Da gab's ne kurze Diskussion drüber, aber da waren nur ein, zwei Leute [...], die das ein bisschen anders gesehen haben und die Prioritäten anders aufgezoogen hätten, aber da habe ich mich dann durchgesetzt.“ (Müller, Beta-D, 135)

Herr Müller kommt später im Gespräch noch einmal auf das Verhältnis von Politikberatung und Grundlagenforschung zurück:

„Da ging's wirklich um die Ausrichtung des Institutes und da gab's eben die eine Fraktion, die sagte,[...] wir haben jetzt genug Grundlagenforschung gemacht, wir wissen, wo wir da stehen, das tragen wir jetzt an gezielten, an ein paar Stellen noch weiter, aber jetzt ist die Mission Politikberatung auf Teufel komm raus. [...] Und die andere Fraktion sagte, wenn wir in die Richtung gehen in aller Konsequenz, dann werden wir wie Institut XY, das es schon gibt, dann treten wir auf dieselbe Ebene und halten nicht die Stärken aufrecht, die das Beta-D groß gemacht haben. [...] Und das war halt ne längere Diskussion, die ging über mehrere Monate hin und her und da hat

*dann irgendwann der Chef ein Machtwort gesprochen und gesagt, so, das ist es jetzt.“
(Müller, Beta-D, 187)*

Die Darstellung von Herrn Müller macht Mehreres deutlich. Zum einen geht es in der Vorbereitung auf die Evaluation ganz offensichtlich nicht allein um gut formulierte Texte in professionellem Layout. Würde man eine reine Wettbewerbssituation annehmen, bei der es allein um das Evaluationsergebnis geht, wäre mehr Taktik durchaus rational. Aber die Evaluation wird gewissermaßen in das Institut „hineingelassen“ und dies, weil die Situation zugleich im Sinne der Normen der Fachgemeinschaft gedeutet wird. Das Gutachtergremium ist zugleich eine Gruppe von Kolleginnen und Kollegen, von denen man eine konstruktive Rückmeldung erwartet.

Deshalb werden in der Vorbereitung Grundverständnisse sorgfältig hinterfragt. Und es werden Entscheidungen getroffen. Das Stichtagsereignis Evaluation zwingt zur Entscheidung über Fragen, die die Arbeit am Institut möglicherweise stets latent begleiten. In diesem Prozess gewinnt der Koordinationsmechanismus Organisation und Hierarchie zwangsläufig an Bedeutung. Wer kann solche Fragen zu einem Abschluss bringen? 25 „ausgewählte“ Personen waren an dem Workshop beteiligt, „da habe ich mich durchgesetzt“ und „da hat der Chef ein Machtwort gesprochen“ – so beschreibt Herr Müller die wesentlichen Instanzen. Im Unterschied zu der als wissenschaftstypisch angenommenen Koordination in einer Kombination aus Konkurrenz und Kooperation geht es bei der Evaluationsvorbereitung um Zuständigkeiten und Kompetenzen, um klassische Organisationsmechanismen also.

Dies ist ein Beispiel; im Vergleich zu den anderen untersuchten Instituten wird für Beta-D an mehreren Stellen eine prinzipiell eher hierarchische Struktur und eine besondere Abhängigkeit vom Institutsleiter beschrieben. Für alle Institute wird aber deutlich, dass mit dem Stichtag Evaluation ein Zwang verbunden ist, sich als Institut mit einer gemeinsam geteilten Orientierung zu präsentieren. Für Beta-D stand hierfür eine Entscheidung zwischen Grundlagenforschung und Politikberatung an, in Alpha-D ging es überhaupt um ein Selbstverständnis als gemeinsames Institut (anstatt individueller Forschung unter einem Dach), bei Gamma-D, Delta-D und Epsilon-D wurde die Organisationsstruktur mit dem Ziel verändert, größere Kohärenz zu erreichen und vor allem darstellbar zu machen. Die damit verbundenen Entscheidungen sind zunächst für die Präsentation und die Unterlagen relevant. Sie sind aber in Zukunft nicht mehr ohne Rechtfertigung zu ignorieren. Die Gutachter beziehen ihre Stellungnahme auf die präsentierten Entscheidungen, und mithin wird dies zum Ausgangspunkt und Maßstab für die nächste Evaluation. Im Falle des Beta-D wurde das für die Evaluation festgeschriebene Selbstverständnis des Instituts durch die Begutachtung bestätigt, dies auch und besonders eine Bestätigung für die Leitungsebene des Instituts. Der durch die Evaluationsvorbereitung gestärkte Koordinationsmechanismus Hierarchie und

Organisation, so jedenfalls liegt es anhand der Darstellungen nahe, dürfte auch über den unmittelbaren Evaluationsprozess hinaus gestärkt bleiben.

5 Die Research Assessment Exercise

5.1 Verfahrensregeln

Die Research Assessment Exercise (RAE) wurde im Jahr 2008 zum sechsten und unter diesem Titel letzten Mal durchgeführt. Eingeführt im Jahre 1986, wurde das Verfahren seitdem kontinuierlich weiterentwickelt (vgl. auch *Martin/Whitley 2010*).⁵ Auf der Grundlage der Ergebnisse werden die institutionellen Mittel für die Forschung an alle Forschungseinrichtungen des Landes vergeben. Zuständig für die Vergabe dieser Mittel ist der Higher Education Funding Council for England (HEFCE) beziehungsweise die entsprechenden Organisationen in Wales, Schottland und Irland. Insbesondere für Universitäten mit einer ausgeprägten Forschungstradition haben die Ergebnisse der RAE finanziell eine hohe Relevanz (*HEFCE 2007*).

Bewertet wird nach Fachgebieten getrennt in so genannten Unit of Assessments (UoA). Im Jahr 2008 wurden 67 UoAs definiert. Universitäten und sonstige Forschungseinrichtungen reichen Bewerbungsunterlagen an die jeweilige UoA, in der neben Daten zur Institution auch jeweils vier Publikationen von allen „forschungsaktiven Wissenschaftlern“ enthalten sind. Alle Unterlagen werden im Laufe eines Jahres in so genannten sub-panels bewertet, Mitglieder sind weit überwiegend Wissenschaftlerinnen und Wissenschaftler aus Großbritannien. Die eingereichten Artikel werden gesichtet, individuell bewertet und mit einer Note von „unclassified“ über 1* (national interessanter Beitrag) bis 4* (weltweit bahnbrechender Beitrag) versehen. Als Ergebnis des Bewertungsverfahrens wird für jedes Institut ein Profil erstellt, das Auskunft darüber gibt, wie viel Prozent der eingereichten Beiträge mit welcher Note versehen wurde. Das Ergebnisprofil wird unmittelbar in eine Formel übertragen, die das Budget des jeweiligen Instituts errechnet. Diese Formel ist allerdings während des Bewertungsverfahrens noch nicht öffentlich; auch die Panelmitglieder wissen nicht, welche Note wie viel genau „wert ist“.

Die Verfahrensregeln der RAE setzen also in großem Maß auf den Wettbewerb als Koordinationsmechanismus, die Wettbewerber konkurrieren unmittelbar um das Erreichen des gleichen Zieles. Bei der RAE handelt es sich dem Grunde nach um einen diszipliniert organisierten Wettbewerb zwischen allen Forschungseinrichtungen des Landes, auf dem Prüfstand steht die finanzielle Grundausstattung, und ein schlechtes

⁵ Aktuell wird unter dem Titel Research Excellence Framework (REF) ein neues Bewertungsverfahren vorbereitet, das 2014 umgesetzt wird. Nachdem zwischenzeitlich auch eine Ablösung des Peer Review-Prinzips durch rein numerische Verfahren in der Diskussion war, bleiben nun wesentliche Bestandteile der RAE auch im REF erhalten. Die Anzahl der Bewertungsgremien wurde reduziert, und der Impact der Forschungsleistung wird zusätzlich bewertet (vgl. *HEFCE 2011*).

Abschneiden führt zu unmittelbaren personellen Konsequenzen. Dabei wird mit jedem Wettbewerb die Ausgangssituation wieder „auf Null“ gestellt, Bestände werden nicht fortgeschrieben. Aus den Vorjahren weiß man zur Zeit des Verfahrens auch, dass nur für die Anteile, die mindestens im oberen Mittelfeld der Notenskala bewertet werden, überhaupt Mittel vergeben werden. Der Wettbewerb sieht außerdem einen festen Topf je Fachgebiet vor. Für ein Institut ist also nicht allein von Bedeutung, dass das eigene Institut möglichst gut abschneidet, sondern von unmittelbar finanziellem Vorteil wäre auch, wenn die anderen möglichst schlecht abschneiden.

5.2 Antizipation in den Instituten

Mit dieser Anlage des Verfahrens wäre zunächst ein wettbewerblich-rationales Verhalten der Beteiligten zu erwarten. Seitens der Institute sind eine Vielzahl von Selektionsentscheidungen gefragt: Die Auswahl der UoA, die Auswahl der forschungsaktiven Wissenschaftler, die Auswahl der Publikationen. Alle diese konkreten Entscheidungen wie auch die Formulierung der Unterlagen sind danach zu orientieren, wie das Ergebnis in der Formel gewichtet wird und welche Entscheidung einen Vorteil gegenüber den Wettbewerbern verspricht. Gefragt ist also eine gute Taktik.

Zugleich basiert auch die RAE auf einem Peer Review. Das Gegenüber für die sich bewerbenden Institute sind die Fachkollegen aus der eigenen wissenschaftlichen Community, nicht etwa eine extern beauftragte Rating-Agentur. Die Frage bleibt also, welcher Mechanismus inwiefern handlungskordinierend wirkt: Inwiefern verstehen die Institutsvertreterinnen und -vertreter den Rahmen als extern initiierten Wettbewerb und nehmen die Logik des Vergleichs und der Taktik an? Wo und inwiefern reibt sich diese Logik mit ihrem Verständnis als Mitglied der Fachgemeinschaft?

Unsere empirische Analyse zeigt, wie unterschiedlich generell die Strategien bei der Erstellung der Bewerbungsunterlagen sind. Insbesondere an der Auswahl der forschungsaktiven Wissenschaftler lässt sich ein Großteil der Haltung der befragten Person zum Verfahren ablesen. In drei von uns untersuchten Instituten finden wir seitens der Führungspersonen drei unterschiedliche Grundhaltungen⁶:

- Herr Gomati (Epsilon-UK) steht dem Verfahren prinzipiell skeptisch gegenüber, er fühlt sich gezwungen, Dinge zu tun, die ihm als Wissenschaftler nicht entsprechen.
- Herr Treßl (Gamma-UK) hält die RAE prinzipiell für ein Verfahren, das auch der eigenen Selbstvergewisserung und Weiterentwicklung dienen kann.

⁶ Dargestellt werden hier die drei Institute, für die eine vollständige Fallstudie im Sinne der in Abschnitt 2 dargelegten Prozesslogik durchgeführt werden konnte. Die wenigen Interviews, die für das vierte Institut (Alpha-UK) zur Verfügung stehen, machen allerdings deutlich, dass hier ebenfalls alle Mitarbeiter einbezogen wurden. Um das Ergebnis zu optimieren, wurden in diesem Institut bereits vier Jahre vor der RAE erhebliche Umstrukturierungen vorgenommen.

- Herr Fredo (Beta-UK) sieht in der RAE einen existenziellen Wettbewerb und nimmt ihn an.

An den drei Instituten wurde mit der Frage, wessen Arbeit Eingang in die Bewerbung findet, unterschiedlich umgegangen. Herr Gomati etwa beschreibt: *“Okay, well the first part is always quiet easy because of Epsilon-UKs ambition and I hope reputation, the ambition is for all academic staff to be submitted in RAE.”* (Gomati, Epsilon-UK, 322)

Für die letzte RAE in 2001 beschreibt Herr Gomati, dass an seinem Institut mehr als 90 Prozent der Mitarbeiter einbezogen waren, und: *“[...] so we actually submitted a higher proportion than whichever one it was, Oxford or Cambridge for example, I mean as a university in fact, because we weren't selective, sorry let's get it right, we didn't game play in the way that some institutions did, and out of Oxford and Cambridge I know that one did and one didn't.”* (Gomati, Epsilon-UK, 326)

Einen möglichst hohen Anteil der Belegschaft als forschungsaktiv zu melden entspricht nach Gomatis Vorstellung der Reputation seines Hauses. Auch Gomati sieht damit das eigene Institut im Wettbewerb, er zieht den Vergleich zu Oxford und Cambridge, um deutlich zu machen, wie gut das eigene Institut in dieser Kennzahl dasteht. Der Vergleich, den Gomati anstellt, ist aber an Kriterien der Wissenschaftsgemeinschaft orientiert, nicht an denen der RAE-Wettbewerbsorganisation. Weil es der Reputation des Instituts abträglich wäre, hält er es für unmöglich, nur einen geringen Anteil der Belegschaft zu melden. Der Gegenbegriff zu seiner der Reputation verpflichteten Strategie ist der des „game playing“. Der Begriff wird später im Interview noch öfter verwendet als etwas, was ihm als Wissenschaftler zuwider ist, wozu er sich aber zum Teil durch die RAE gezwungen sieht. Würde Gomati ganz in den Kategorien des RAE-Wettbewerbs denken, dann ist diese Berufung auf Reputation nicht rational, denn zum Zeitpunkt der Bewerbung war bereits bekannt, dass eine inklusive Strategie der Personalauswahl keine positiven Auswirkungen auf das Budget haben würde.

Auch am Institut, das Treßl vertritt, wurden nahezu alle Wissenschaftler mit ihren Publikationen gemeldet. Die Erklärung der Strategie allerdings fällt hier etwas anders aus: *“We put everybody in. [...] in 2001 and 1996 and 1992 [...], two or three people were left out. And it's quite divisive. Okay, so I was in an easier position, but I did have one member of staff, okay, who I couldn't have returned because he had nothing, he had no outputs, no papers. Okay, but I solved his problem by a sneaky method two, which is, he's ill, so I got him to retire on ill health grounds there, otherwise I would have been, but if he hadn't I would have put him in. It was more important to me to have one passenger than to have somebody who is left out because it's quite divisive.”* (Treßl, Gamma-UK, 31)

Treßl wägt in seiner Strategie die Innenwirkung ins Institut gegen die potenziellen Wettbewerbsvorteile ab. Er sieht sich in der glücklichen Situation, niemanden explizit aus inhaltlichen Gründen ausschließen zu müssen. Stünde er aber vor einer solchen Entscheidung, so würde er den Schaden, den eine Nicht-Berücksichtigung eines einzelnen Wissenschaftlers im Institut auslösen würde, größer bewerten als den Gewinn, der im RAE-Verfahren damit zu erzielen wäre. Die prinzipielle Wettbewerbslogik des Verfahrens ist damit Teil seiner Überlegungen. Anders als Gomati sorgt sich Treßl nicht in erster Linie um die Reputation seines Instituts beziehungsweise sieht diese nicht durch eine niedrige Quote der gemeldeten Wissenschaftler gefährdet, vielmehr kalkuliert er innerhalb der Wettbewerbsregeln. Im weiteren Verlauf betont er, dass er zwar einen „passenger“ zum Wohle des Institutsklimas in die Bewerbung integrieren würde. Bei mehreren allerdings müsste er abwägen, es könnte sein, dass dann der Wettbewerbsnachteil zu groß würde. Treßl macht deutlich, wie er die verschiedenen Entscheidungen taktisch abgewogen hat – etwa die Frage, ob die Arbeiten aller Wissenschaftler für eine UoA eingereicht werden oder ob man die Gruppe teilen sollte. Er beschreibt den gesamten Vorgang der Bewerbungsvorbereitung ähnlich einem Strategiespiel und kalkuliert Risiken und Gewinnmöglichkeiten. Damit trennt er Wissenschaft und deren Bewertung vom Prozess der RAE:

„because the research assessment exercise is not about assessing research, it's about getting money in the UK. [...] So actually, I mean there are some tactical decisions not to do with how good the research is at all, it's purely about what comes afterwards. So it's all a bit, it's quite tricky.“ (Treßl, Gamma-UK, 83)

Am Institut, das Treßl vertritt, wird also prinzipiell die Wettbewerbslogik im Sinne eines Spiels aufgenommen, an dessen Regeln man sich orientiert. Ein Kollege von Treßl berichtet in ähnlicher Weise, wie er bei der Auswahl der Publikationen zwischen den Bewertungsmaßstäben der scientific community und denen des panels unterscheidet.

Am Institut, das Fredo vertritt, wurden bereits zu Beginn des Jahres 2005 diejenigen Mitarbeiterinnen und Mitarbeiter ausgewählt, deren Publikationen in die Bewerbung aufgenommen werden sollten. Dafür wurde eine RAE für die Universität simuliert, alle Wissenschaftlerinnen und Wissenschaftler wurden aufgefordert, ihre besten Publikationen zu diesem Zeitpunkt auszuwählen, externe und interne Gutachter wurden bestellt und in fachlich zusammen gesetzten Teams wurden alle Publikationen bewertet. Fredo definiert die eigene Universität stets im Wettbewerb mit den anderen, und er ist bemüht, sich Wettbewerbsvorteile zu verschaffen, indem er beispielsweise diesen Bewertungsprozess früher organisierte als die meisten anderen.

Die Bewerbungsstrategie des Instituts war es, nur diejenigen einzubeziehen, für deren Publikationen als Durchschnitt mindestens eine 3 erwartet wurde: *“because my ap-*

proach was since this is going to determine our research reputation, in a way that no other RAE had determined our research reputation, then I wanted to make sure that our research halo shone brightly.” (Fredo, Beta-UK, 65)

Anders als Gomati setzt Fredo die Reputation des Institutes mit dem Ergebnis der RAE gleich, unabhängig davon, welchen Anteil von Wissenschaftlern er einbezieht. Und anders als Treßl handelt es sich aus seiner Sicht auch hier nicht in erster Linie um einen Prozess zur Verteilung von Geld, der sich von dem Wettbewerb um Reputation an manchen Stellen trennen lassen würde. Die aktuelle RAE schätzt er dabei als besonders bedeutsam für die Reputation des Institutes ein, weil er von einem Wechsel des Bewertungsverfahrens ausgeht und dann die Ergebnisse der letzten RAE um so länger im Gedächtnis bleiben würden.

Fredo übernimmt die Kategorien und die Logik der RAE als Ziele und Leitideen seiner Institutsführung. Der zentrale Hebel, um sich im Wettbewerb Vorteile zu verschaffen, sind Publikationen in „guten“ Zeitschriften. Entsprechend sollten die Jahre nach dem internen Auswahlverfahren dazu genutzt werden, den Anteil der Artikel in den Zeitschriften zu steigern, die auf dem Zitationsindex weit oben stehen.

Die Mitarbeiter des Instituts werden seit dem internen Auswahlverfahren in zwei Gruppen unterteilt. Diejenigen, deren Publikationen in die Bewerbung aufgenommen werden sollen, sollen in der Forschung und insbesondere in der guten Platzierung der Artikel unterstützt werden. Diejenigen, die nicht ausgewählt wurden, sollen sich verstärkt „anderen“ Aufgaben, in erster Linie der Lehre widmen, um damit auch die forschungsaktiven Wissenschaftler zu entlasten. Über diejenigen, die nicht einbezogen werden, sagt Fredo:

“On advice from HR, we also came to the agreement that, it didn’t necessarily mean their research career was at an end, what it meant was after the RAE [...] that the people who were not going to be submitted would be invited to submit a research plan to their head of school, and the research plan would be discussed with their head of school. If the research plan was accepted, then we would support those people to become research active again. If the research plan was not accepted, then they would be asked to continue with a double load of teaching, more administration, they wouldn’t be research active. And then there would be a change in the terms and conditions.” (Fredo, Beta-UK, 171)

Die RAE gibt hier die Orientierung für die Personalentwicklung. Die Logik des Wettbewerbs in der konkreten Organisationsform, wie sie in der RAE angelegt ist, wird damit zur Grundlage der Organisation der Forschung am Institut. Übernommen werden das Prinzip, Artikel zu einem Stichtag in vier Noten zu bewerten ebenso wie die Defi-

dition der Noten und der Prozess der Bewertung. Das Ergebnis entscheidet damit über künftige Karrierechancen.

Während Gomati die Prinzipien der Wissenschaft als zum Teil unvereinbar mit denen der RAE beschreibt und Treßl die RAE als einen Prozess ansieht, an dem man sich beteiligen kann, ohne die eigenen wissenschaftlichen Prinzipien zu gefährden, beschreibt Fredo die wissenschaftlichen Prinzipien an seinem Institut als deckungsgleich mit denen der RAE. Anders ausgedrückt: Für Fredo besteht, anders als in unterschiedlicher Abstufung für Gomati und Treßl, kein Konflikt zwischen den wissenschaftsintern akzeptierten Koordinationsmechanismen und dem durch die RAE organisierten Wettbewerb.

Die Darstellungen machen deutlich, dass die RAE durchaus geeignet ist, Wettbewerb als Koordinationsmechanismus in der Wissenschaft zu stärken. Außerdem zeigen die drei Beschreibungen, dass die Verfahrensregeln allein noch nicht darüber entscheiden, inwiefern der etablierte Koordinationsmechanismus in Frage gestellt wird. Gomati und Treßl unterscheiden jedenfalls zwischen den Anforderungen der RAE und den von Ihnen als wissenschaftsadäquat angenommenen Regeln. Anzunehmen ist, dass der Einfluss der RAE hier auf die institutsinternen Koordinationsmechanismen weniger substantiell ist als am Institut von Herrn Fredo. Wovon genau es abhängt, wie umfassend der in der RAE angelegte Wettbewerbsmechanismus angenommen wird, lässt sich anhand unseres Materials nicht nachvollziehen. Disziplinäre Faktoren kämen als Erklärung ebenso in Frage wie die Organisationsstruktur des Instituts und die Persönlichkeit der Leitung. Für die Fragestellung dieses Artikels soll zunächst herausgestellt werden, dass auch ein stringent auf Wettbewerbsmechanismen zielendes Verfahren wie die RAE nicht ungebrochen Eingang in die wissenschaftsinterne Koordinationslogik findet – die Normen der Fachgemeinschaft bleiben Teil des allerdings neu kombinierten Mechanismus.

Schließlich lässt sich wie schon für die Verfahren der Leibniz-Gemeinschaft auch für die RAE feststellen, dass hierarchische Entscheidungen an Bedeutung gewinnen. Werden im Falle der Leibniz-Gemeinschaft abschließende Entscheidungen über Institutsstrukturen erforderlich, so geht es im Falle der RAE um zum Teil existenzielle Fragen der Personalentwicklung. Nicht an einer Bewerbung beteiligt zu sein, kann für Einzelne das Karriereaus bedeuten. Damit erhalten Personalverantwortliche an den Instituten einen Einfluss und eben auch eine Verantwortung, wie sie üblicherweise für Wissenschaftsorganisationen nicht angenommen werden.

6 Fazit

Inwiefern können Evaluationen, deren Ergebnisse zugleich Grundlage für Allokationsentscheidungen sind, die wissenschaftsinternen Koordinationsmechanismen verändern? Mit dieser Fragestellung widmete sich der Beitrag einer Untersuchung von Evaluationsverfahren in Deutschland (Leibniz-Gemeinschaft) und Großbritannien (RAE). Als in der Wissenschaft dominierende Koordinationsmechanismen wurde eine Kombination von Wettbewerb (Konkurrenz um die besten Ideen) und Gemeinschaft (Kooperation unter Fachkollegen) eingeführt, während Hierarchie und Organisation in der Wissenschaft traditionell als wenig bedeutsam angenommen wird.

Im Ergebnis zeigt sich zunächst, dass die mit den Verfahren jeweils verbundenen Mechanismen durchaus ihre Wirkung in die Institute hinein zeigen und plausibel ist, dass diese Wirkungen Bedeutung über den konkreten Evaluationstermin hinaus haben. Im Falle der Leibniz-Gemeinschaft wird ein Wettbewerb zwischen den Instituten organisiert, die Evaluierten sind qua Regeln aufgefordert, ihre Besonderheit so überzeugend dazustellen, dass sie damit ihre weitere Förderfähigkeit begründen. Im Falle der RAE konkurrieren letztlich die jeweils in Forschungseinheiten zusammengefassten Individuen miteinander und es gilt, einen besseren Output zu präsentieren als die anderen Einheiten.

Beide Verfahren lösen erhebliche Aktivitäten aus, von denen ein erkennbarer Anteil als strategische Aktivitäten im Sinne der Wettbewerbsregeln zu identifizieren ist. Allerdings zeigt sich gleichzeitig für beide Verfahren, dass diese Wettbewerbsregeln nicht ungebrochen angenommen werden, dass strategisches Verhalten vielmehr neu kombiniert wird mit den Koordinationsmechanismen der Fachgemeinschaft. Im Falle der Leibniz-Gemeinschaft bedeutet das, dass die Evaluierung als zwar erzwungene, aber dann gewissenhaft angenommene Unterstützung der eigenen Organisationsentwicklung antizipiert wird. Deutlich intensiver als dies allein aufgrund der Regeln notwendig wäre, werden Institute in ihrem Selbstverständnis neu hinterfragt, werden Strukturen und inhaltliche Ausrichtungen verändert. Im Falle der RAE zeigt sich ein differenziertes Bild. Die Regeln legen strategisches Kalkül nahe – konkret weiß man zum Zeitpunkt der Bewerbung, dass es finanziell vorteilhaft wäre, nur die besten Wissenschaftler zu melden. In der Kombination mit den Mechanismen der Fachgemeinschaft wird dies aber in einem der drei untersuchten Institute als der eigenen Reputation nicht angemessenes „game playing“ abgelehnt, im anderen Fall wägt der Institutsleiter den internen Schaden gegen finanzielle Chancen ab, und der dritte lässt sich auf das Kalkül ein und macht es auch zur Maßgabe der weiteren Personalpolitik. Selbst im stringenten Regime der RAE können also die Mechanismen der Fachgemeinschaft nicht ohne weiteres ausgehebelt werden. Evaluationen, auch wenn mit ihnen existenzielle Entscheidungen zur Mittelallokation verbunden sind, führen nicht per se zu mehr oder zu der Wissenschaft bis dato fremden Formen des Wettbewerbs.

Deutlich geworden ist für beide Verfahren, dass der ansonsten in der Wissenschaft als wenig bedeutsam angenommene Koordinationsmechanismus von Organisation und Hierarchie erheblich gestärkt wird. Evaluationen erfordern Entscheidungen zu einem Stichtag – Entscheidungen über Fragen, die ansonsten möglicherweise nie entschieden, sondern als laufende Auseinandersetzungen ein Institut stets begleiten würden. Eine Aufwertung der Organisation in diesem Sinne kann auch durchaus als wissenschaftspolitisch intendiertes Ziel angenommen werden. In dem Maße wie Wissenschaft und Forschung als Standortfaktor im internationalen Wettbewerb angenommen werden, gewinnen auch Begriffe wie Profilbildung und Alleinstellungsmerkmal an Bedeutung. Aus der Perspektive der Wissenschaftsforschung allerdings lassen diese Entwicklungen aufmerken, weil eine konzeptionelle Orientierung, was eine „starke Wissenschaftsorganisation“ ausmacht, weithin zu fehlen scheint. Wenn die bislang viel zitierten Konzepte der organisierten Anarchie oder der losen Kopplung sich unter aktuellen wissenschaftspolitischen Rahmenbedingungen nicht weiter bewähren können, und wenn man gleichzeitig die unreflektierte Übernahme von diversen Moden der betriebswirtschaftlich orientierten Organisationstheorie vermeiden will, dann ist eine neue theoretisch und empirisch fundierte Auseinandersetzung mit den Organisationsbedingungen für „gute Wissenschaft“ dringend erforderlich.

Literatur

Benz, Arthur (2007): Politischer Wettbewerb. In: Benz, Arthur u. a. (Hrsg.): Handbuch Governance: Theoretische Grundlagen und empirische Anwendungsfelder. Wiesbaden, S. 54–67

Bourdieu, Pierre (1975): Specificity of Scientific Field and Social Conditions of Progress of Reason. In: Social Science Information 14, 1975, 6, S. 19–47

Brunsson, Nils; Sahlin-Andersson, Kerstin (2000): Constructing Organizations: The Example of Public Sector Reform. In: Organization Studies 21, 2000, S. 721–746

Cohen, Michael D.; March, James G.; Olsen, Johan P. (1972): A Garbage Can Model of Organizational Choice. In: Administrative Science Quarterly 1/1972, 1972, S. 1–25

Döhler, Marian (2007): Hierarchie. In: Benz, Arthur u. a. (Hrsg.): Handbuch Governance: Theoretische Grundlagen und empirische Anwendungsfelder. Wiesbaden, S. 46–53

Frey, Bruno S. (2008): Evalutis – eine neue Krankheit. In: Matthies, Hildegard u. a. (Hrsg.): Wissenschaft unter Beobachtung. Effekte und Defekte von Evaluationen. Wiesbaden, S. 125–140

Gläser, Jochen (2007): Gemeinschaft. In: Benz, Arthur u. a. (Hrsg.): Handbuch Governance. Wiesbaden, S. 82–92

Gülker, Silke; Simon, Dagmar (2011): Nach der Evaluation ist vor der Evaluation. Institutionelle Folgen von Forschungsbewertungen im internationalen Vergleich. In: Horn-

bostel, Stefan u. a. (Hrsg.): Evaluation: New Balance of Power? iFQ Working Paper No. 9. Dezember 2011. Berlin, S. 93–93

Gülker, Silke; Simon, Dagmar; Torka, Marc (2012): Evaluation of Science as Consultancy. In: Quaderni Special issue: Research Evaluation: Reopening the Controversy, 2012, No. 77 2011/2012, S. 41–54

Hagstrom, Warren O. (1974): Competition in Science. In: American Sociological Review 39/1974, 1974, S. 1–18

Hayek, Friedrich A (1969): Der Wettbewerb als Entdeckungsverfahren. In: Hayek, Friedrich A (Hrsg.): Freiburger Studien. Tübingen, S. 249–265

HEFCE, Higher Education Funding Council (2011): REF 2014. Research Excellence Framework. Assessment Framework and Guidance on Submissions. http://www.hefce.ac.uk/research/ref/pubs/2011/02_11/02_11.pdf (Zugriff: 21. Dezember 2011)

HEFCE, Higher Education Funding Council for England (2007): Funding Higher education in England. How HFECE allocates its funds. Guide July 2007/20. Bristol

Kieser, Alfred (2010): Unternehmen Wissenschaft? In: Leviathan 3/2010, 2010, S. 347–367

Kromrey, Helmut (2003): Qualität und Evaluation im System Hochschule. In: Stockmann, Reinhard (Hrsg.): Evaluationsforschung. Opladen, S. 233–258

Lamont, Michèle (2009): How Professors Think: Inside the Curious World of Academic Judgment. Cambridge

Latour, Bruno; Woolgar, Steve (1986): Laboratory Life. The Construction of Scientific Facts, 2 Edition. Princeton (NJ)

Leibniz-Gemeinschaft (2007): Kriterien für die Evaluierung von Einrichtungen der Leibniz-Gemeinschaft (SEN 0057/07, 18.07.2007) Im Internet: <http://www.leibniz-gemeinschaft.de/?nid=evqu&nidap=&print=0>, (Zugriff: 05. Oktober 2008)

Lüders, Christian; Meuser, Michael (1997): Deutungsmusteranalyse. In: Hitzler, Ronald u. a. (Hrsg.): Sozialwissenschaftliche Hermeneutik. Opladen, S. 57–79

Martin, Ben; Whitley, Richard (2010): The UK Research Assessment Exercise: A Case of Regulatory Capture? In: Whitley, Richard u. a. (Hrsg.): Reconfiguring Knowledge Production. Changing Authority Relationships in the Science and their Consequences for Intellectual Innovation. Oxford, S. 51–80

Matthies, Hildegard; Matthäus, Sandra (2010): Science Between Organization and Profession: Opportunities for Neutralizing Gender? – Reflections on Research Assessment. In: Riegraf, Birgit u. a. (Hrsg.): GenderChange in Academia. Re-Mapping the Fields of Work, Knowledge, and Politics from a Gender Perspective. Wiesbaden, S. 87–100

Merton, Robert K. (1973): The Normative Structure of Science. In: Storer, N.W. (Hrsg.): The Sociology of Science: Theoretical and Empirical Investigations. Chicago, IL, S. 267–278

Mintzberg, Henry (1979): The Structuring of Organizations. A Synthesis of the Research. Englewood

Neidhardt, Friedhelm (1988): Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG. Opladen

Popper, Karl R. (1959): The logic of scientific discovery. New York

Power, Michael (2008): Research Evaluation in the Audit Society. In: Matthies, Hildegard u. a. (Hrsg.): Wissenschaft unter Beobachtung-Effekte und Defekte von Evaluation. Wiesbaden, S. 15–24

Soeffner, Hans-Georg (1989): Prämissen einer sozialwissenschaftlichen Hermeneutik. In: Soeffner, Hans-Georg: Auslegung des Alltags. Der Alltag der Auslegung. Zur wissenssoziologischen Konzeption einer sozialwissenschaftlichen Hermeneutik. Frankfurt am Main, S. 66–97

Torka, Marc (2011a): Der Evaluator. Positionierungen, Strategien, Deutungs-, Handlungs- und Urteilsweisen von Wissenschaftlern in Prozessen der Wissenschaftsevaluation. In: Sozialer Sinn – Zeitschrift für hermeneutische Sozialforschung 1, 2011, S. 115–138

Torka, Marc (2011b): Institutioneller gleich handlungspraktischer Wandel? Das Beispiel von Begutachtungspraktiken bei der Evaluation wissenschaftlicher Einrichtungen. In: Hornbostel, Stefan u. a. (Hrsg.): Evaluation: New Balance of Power? iFQ Working Paper No. 9. Dezember 2011. Berlin, S. 69–81

Weber, Max (1980 [1922]): Wirtschaft und Gesellschaft. Tübingen

Wernet, Andreas (2006): Einführung in die Interpretationstechnik der Objektiven Hermeneutik. In: Bohnsack, Rolf u. a. (Hrsg.): Qualitative Sozialforschung, Band 11. Wiesbaden, S.

Whitley, Richard (2010): Reconfiguring the Public Science. The Impact of Governance Changes on Authority and Innovations in Public Science Systems. In: Whitley, Richard u. a. (Hrsg.): Reconfiguring Knowledge Production. Changing Authority Relationships in the Science and their Consequences for Intellectual Innovation. Oxford, S. 3–47

Wiesenthal, Helmut (2000): Markt, Organisation und Gemeinschaft als "zweitbeste" Verfahren sozialer Koordination. In: Werle, Raymond u. a. (Hrsg.): Gesellschaftliche Komplexität und kollektive Handlungsfähigkeit. Frankfurt am Main; New York, S. 44–73

Anschrift der Verfasserin:

Dr. Silke Gülker

WZB Wissenschaftszentrum Berlin für Sozialforschung

Forschungsgruppe Wissenschaftspolitik

Reichpietschufer 50

10785 Berlin

E-Mail: guelker@wzb.eu

Das sächsische Hochschulrating als Beispiel für einen evidenzbasierten Ansatz in der Hochschulentwicklungsplanung – Ein Werkstattbericht

Ulf Banscherus, Mirjam Reiß

Mit dem sächsischen Hochschulrating wurde im Freistaat Sachsen ein Modellprojekt durchgeführt, das die staatliche Hochschulentwicklungsplanung vor dem Hintergrund des demographischen Wandels auf eine evidenzbasierte Grundlage stellen sollte. Hierzu wurde ein mehrstufiges Bewertungsverfahren ausgewählter Fächer an den Universitäten und Fachhochschulen entwickelt, das eine umfangreiche quantitative Datenauswertung mit einer qualitativen Betrachtung durch externe Expertinnen und Experten verband. Die Ergebnisse, die übergreifende strategische Empfehlungen an die Hochschulpolitik genauso beinhalten wie Vorschläge zur stärkeren Abstimmung des Fächerangebots zwischen den Hochschulen bzw. zur Konzentration des Studienangebots auf bestimmte Standorte, wurden aufgrund der hohen Transparenz des Verfahrens und der umfangreichen Beteiligungsmöglichkeiten von den Hochschulen weitgehend anerkannt und bildeten die Basis für die Fortschreibung des sächsischen Hochschulentwicklungsplans.

1 Einleitung

Die staatliche Hochschulentwicklungsplanung steht bereits seit längerem vor einer doppelten Herausforderung: Auf der einen Seite erwarten Politik und Öffentlichkeit von den Wissenschaftsministerien, dass diese auf aktuelle Entwicklungen wie beispielsweise die doppelten Abiturjahrgänge reagieren und eine ausreichende Zahl von Studienplätzen an den Hochschulen sicherstellen. Sie sollen ebenfalls Sorge dafür tragen, dass die Hochschulen mit öffentlichen Mitteln effizient umgehen und diese effektiv einsetzen. Auf der anderen Seite verwahren sich die Hochschulen gegenüber Eingriffen in die ihnen zugesicherte Autonomie und finden hierfür durchaus Unterstützung bei politischen Parteien und Verbänden sowie in der Medienberichterstattung. Erschwerend kommt hinzu, dass die Ministerien häufig gar nicht über die notwendigen Informationen verfügen, um Entscheidungen in Detailfragen treffen zu können.

In diesem Beitrag soll mit dem sächsischen Hochschulrating ein Modellprojekt vorgestellt werden, das die Zielsetzung verfolgte, die Hochschulentwicklungsplanung

des Freistaates Sachsen auf eine evidenzbasierte Grundlage zu stellen. Die nach wie vor politisch zu treffenden Planungsentscheidungen für die sächsische Hochschullandschaft sollten nicht mehr auf den häufig als subjektiv kritisierten Einschätzungen und Bewertungen der Verwaltung basieren, sondern auf der Grundlage eines empirisch fundierten Entwicklungsszenarios erfolgen, das unter anderem Prognosen und Projektionen zur künftigen Studiennachfrage an den sächsischen Hochschulen und zum regionalen Fachkräftebedarf einbezieht. Ein solches Bewertungsverfahren, das als Grundlage und Legitimation politischer Entscheidungen dienen soll, muss auf klaren Kriterien und einer im Vorfeld formulierten Zielstellung basieren, wissenschaftliche Expertise genauso einbeziehen wie die Sicht der von hochschulpolitischen Entscheidungen potenziell Betroffenen und weiterhin ein hohes Maß an Transparenz aufweisen, um die für weitreichende hochschulpolitische Strukturentscheidungen notwendige Akzeptanz durch Hochschulen und Öffentlichkeit erreichen zu können. Dies stellt die staatliche Hochschulplanung vor eine schwierige, aber gleichwohl nicht unlösbare Aufgabe. Das sächsische Hochschulrating versucht, diesen Anforderungen zumindest nahe zu kommen, indem es ein hochschulplanerisches Instrument darstellt, das eine umfassende indikatorbasierte Analyse von ausgewählten Fächern mit einem qualitativen Bewertungsprozess verbindet, um so eine differenzierte Leistungsanalyse in mehreren Bewertungsdimensionen zu erreichen. Ein Abgleich zwischen dem Trendszenario und der Leistungsanalyse sollte in diesem Zusammenhang die empirische Basis liefern, auf deren Grundlage eine evidenzbasierte Hochschulentwicklungsplanung möglich sein sollte.

In den nachstehenden Abschnitten erfolgen zunächst eine Reflexion der Rahmenbedingungen staatlicher Hochschulentwicklungsplanung im Kontext der seit den 1990er Jahren zu beobachtenden weitgehenden Deregulierung sowie eine Diskussion des Verfahrens des Peer Review einschließlich der damit verbundenen Kritik. Daran anschließend werden Ziele, Vorgehensweise, Indikatoren und Ergebnisse des sächsischen Hochschulratings in der gebotenen Kürze vorgestellt. Abschließend werden wichtige Erfahrungen mit diesem Verfahren diskutiert.

2 Hochschulsteuerung im Wandel

Spätestens seit Mitte der 1990er Jahre haben in Deutschland weitreichende Veränderungen bei der Hochschulsteuerung stattgefunden. So wurden im Rahmen der Stärkung der Hochschulautonomie vielfältige Entscheidungskompetenzen aus der Zuständigkeit der Wissenschaftsministerien der Länder in die Zuständigkeit der Hochschulen, und hier vor allem der Hochschulleitungen beziehungsweise der Hochschulräte, überführt. Dies gilt nicht nur für den Haushaltsvollzug, sondern häufig auch für Strukturentscheidungen wie beispielsweise die Binnenorganisation der Hochschulen oder das Studienangebot. Die Hochschulsteuerung erfolgt zunehmend weniger in

Form von ministerieller Detailsteuerung als vielmehr auf dem Weg von Hochschulverträgen und Zielvereinbarungen, die zwischen dem zuständigen Ministerium und der einzelnen Hochschule abgeschlossen werden. Gestärkt wurden ebenfalls Wettbewerbselemente in der Ressourcenverteilung. Auf diese Weise sollten die Hochschulleitungen in eigener Verantwortung Entwicklungsstrategien formulieren, implementieren und evaluieren können (vgl. *Volter 2007; Zechlin 2007*). Der vertraglich geregelte Steuerung durch Zielvereinbarungen und Hochschulverträge liegt der Gedanke zugrunde, dass die strategischen Entwicklungsziele zunehmend zwischen Bundesländern und Hochschulen als weitgehend gleichberechtigten Partnern ausgehandelt werden sollten, die Umsetzung allerdings überwiegend in der Verantwortung der „autonomen“ Hochschulen liegen sollte – wenn auch mit recht weitgehenden Berichtspflichten gegenüber dem Staat. Problematisch erweist sich in diesem Zusammenhang jedoch, dass Zielvereinbarungen häufig weniger strategische Entwicklungsziele als vielmehr konkrete Maßnahmen umfassen, was die von Politik und Wissenschaft weitgehend erwünschte Deregulierung in nicht unwesentlicher Weise konterkarieren kann (vgl. *Pasternack 2003*).

Der Übertragung neuer Freiheiten, vor allem in finanzieller und organisatorischer Hinsicht, steht aber auch aus Sicht der Befürworterinnen und Befürworter der Hochschulautonomie ein legitimes Interesse an staatlicher Regulierung gegenüber, da „(ein) blindes staatliches Vertrauen in die Weisheit der Hochschulangehörigen beim Umgang mit Steuergeldern“ nicht vorausgesetzt werden könne. Aus diesem Grunde könne auf die staatliche Steuerung auch in einem weitgehend deregulierten Hochschulsystem nicht vollständig verzichtet werden, notwendig sei allerdings eine Neujustierung der Aufgaben der Wissenschaftsministerien (*Ziegele 2002, S. 107*). Dies könne unter anderem durch eine Rückführung der Aktivitäten auf die ministeriellen Kernaufgaben geschehen, wozu ganz wesentlich die Entwicklung einer längerfristigen, hochschulübergreifenden Gesamtstrategie im jeweiligen Bundesland gehört, die aus der „strukturpolitischen Gesamtverantwortung“ des Staates für das Wissenschaftssystem insgesamt resultiere. Eine konkrete landesplanerische und steuerungspolitische Aufgabe sei in diesem Zusammenhang vor allem die Sicherung einer regional ausgewogenen Angebotsstruktur, was auch Fragen der Zusammenlegung oder Kooperation von Standorten umfasse. Hinzu komme die Beantwortung von Standortfragen im Spannungsfeld von Regional- und Wissenschaftspolitik. Entsprechende staatliche Eingriffe in die Autonomie der Hochschulen seien erforderlich, da sich die „Summe der Eigeninteressen (...) nicht automatisch mit dem gesellschaftlich Notwendigen“ decke. Wettbewerb könne unter anderem dazu führen, dass bestimmte Fächer an mehreren Hochschulen angeboten werden, obwohl sie an einzelnen Hochschulen eine unterkritische Größe erreichen oder dass Hochschulen aufgrund der Orientierung am internationalen Forschungswettbewerb die Aufgabe der Versorgung regionaler Arbeitsmärkte mit Fachkräften aus dem Auge verlieren könnten (*Müller u. a. 2010, S. 15f.*).

3 Notwendige Legitimation von Planungsentscheidungen

Vor dem Hintergrund der gesteigerten Hochschulautonomie, nicht zuletzt aber auch aufgrund knapper werdender finanzieller Mittel und sich verändernder gesellschaftlicher Rahmenbedingungen wie dem demographischen Wandel, der insbesondere die ostdeutschen Bundesländer vor große Herausforderungen stellt, steht die Politik vor der Aufgabe, Ressourcen in nachvollziehbarer Weise an die Hochschulen zu vergeben und die Entwicklungsplanung auf einer soliden Informationsbasis vorzunehmen. Das Sächsische Hochschulgesetz sieht hierzu seit der Anfang 2009 in Kraft getretenen Novelle den Aufbau eines übergreifenden Dateninformationssystems vor, das die empirische Grundlage der Hochschulentwicklungsplanung des Freistaates Sachsen bilden soll. Die Verwendung von Indikatoren und Kennzahlen bildet aber keineswegs zwingend kausale Zusammenhänge ab, da diese häufig stark aggregiert sind und so die Inhalte und die Dynamik der Wissenschaft kaum präzise genug abbilden können, um Planungsentscheidungen zu legitimieren. Die Vision einer rein indikatorbasierten Hochschulentwicklungsplanung würde Gefahr laufen, „nichts anderes als ‚tödliche Tabellen‘ zu erzielen“ (Salais 2008, S. 209).

Ein entsprechendes hochschulplanerisches Instrument, das über die etablierten Controlling-Verfahren hinausgeht, sich in der Praxis bewährt hätte und breit akzeptiert wäre, ist derzeit allerdings nicht bekannt. Auch die Übernahme etablierter Verfahren der Qualitätsmessung, Qualitätssicherung und Qualitätsentwicklung von Forschung und Lehre stieße schnell an Grenzen, da diese in aller Regel einen rein wissenschaftsinternen Fokus aufweisen und deshalb für Zwecke der übergreifenden Landesplanung nur bedingt geeignet sind. Als hilfreich kann sich an dieser Stelle die Nutzung eines Ratings erweisen, bei dem die Bewertung der Leistungen der einzelnen Fächer sowie der Hochschulen insgesamt auf den Einschätzungen von Expertinnen und Experten basiert, die die verschiedenen Indikatoren und Kennzahlen sichten, interpretieren und bei der Bewertung weitere (qualitative) Informationen berücksichtigen (Lange 2010, S. 325).

4 Die Rolle von Expertenkommissionen und Peer Review

Öffentliches Handeln benötigt immer auch „Orientierungen und Bezugspunkte, um daran ihren Kurs auszurichten, ihre Zielvorstellungen besser verfolgen und überprüfen zu können und um die Akteure in einen Lernprozess zu versetzen“ (Salais 2008, S. 194). Aus einer hochschulplanerischen Perspektive heraus werden vor diesem Hintergrund neuartige, auf Indikatoren basierende Bewertungsverfahren unter Beteiligung von Wissenschaftlerinnen und Wissenschaftlern benötigt, welche die Anforderung erfüllen, vergleichende Informationen und reflektierte Bewertungen von wissenschaftlichen Leistungen bereitzustellen. Hierbei sollen sie auch Indikatoren wie den Studienerfolg,

die Studiengangauslastung und die Absolventenzahl berücksichtigen. Diese müssen für die staatliche Hochschulentwicklungsplanung als besonders relevant gelten; sie legitimieren auf diese Weise anschließende politische Entscheidungen über die Verteilung finanzieller Mittel und die Struktur des Studienangebots. Hierzu bietet es sich an, im Sinne eines Ratings Elemente des wissenschaftsinternen Ansatzes des Peer Review mit typischen Elementen der Arbeit von Expertenkommissionen zu verbinden.

Bei der Behandlung von Fragen mit einem besonderen Komplexitätsniveau werden seitens der Politik häufig verschiedene Formen der wissenschaftlichen Politikberatung nachgefragt. Von den Regierungen im Bund und in den Ländern werden beispielsweise häufig Expertenkommissionen eingesetzt, um Vorschläge für politische Entscheidungen von besonderer Reichweite vorzubereiten und durch die Einbeziehung wissenschaftlicher und zivilgesellschaftlicher Expertise zu legitimieren. Dies wird auch als Versuch verstanden, aus Informationsdefiziten entstehende staatliche Steuerungsdefizite zumindest teilweise zu kompensieren. Dabei sind die Anforderungen an wissenschaftliche Politikberatung hoch: *„Sie muss die komplexen Problemlagen klar strukturieren und ‚vermessen‘ und somit für den öffentlichen Diskurs klar strukturieren und zudem noch Handlungsmöglichkeiten ausloten, die dann im politischen Raum weiter verarbeitet werden können und auch in die Gesetzgebung einfließen.“* (Heinze 2009, S. 189ff., Zitat: S. 193). Die Inanspruchnahme von wissenschaftlicher Politikberatung ist in nahezu allen Politikfeldern anzutreffen. Dies gilt auch für die Bildungspolitik, wo deren Akzeptanz im Gefolge der politischen „Verarbeitung“ der verschiedenen internationalen Vergleichsstudien wie IGLU und PISA in den vergangenen Jahren deutlich gestiegen ist (Hepp 2011, S. 91ff.). Einer umfassenden Analyse der Arbeit der von der rot-grünen Bundesregierung zwischen 1998 und 2005 eingesetzten Expertenkommissionen zufolge weisen diese einige generalisierbare Merkmale auf. So werden Expertenkommissionen in der Regel für einen begrenzten Zeitraum von einem Bundesministerium innerhalb seiner Ressorthoheit eingesetzt, verfolgen einen klaren Auftrag – zumeist sollen sie Wege aufzeigen, wie mit bereits festgestelltem Veränderungsbedarf umzugehen ist –, und setzen sich aus 13 bis 18 externen Fachleuten zusammen, wobei diese aufgrund ihrer Zusammensetzung zumeist nicht als primär wissenschaftlich orientierte Beratungsgremien gelten können. Berichte von Expertenkommissionen werden in der Regel veröffentlicht, wenngleich hierzu keine formale Verpflichtung besteht (Siefken 2007, S. 314ff.).

Die Einbeziehung des Peer Review-Verfahrens in staatliche Entscheidungsprozesse ist keineswegs unumstritten, da den Kritikerinnen und Kritikern zufolge bei einer solchen Bewertung wissenschaftlicher Güte in hohem Maße implizite Kriterien Anwendung finden, sodass der Staat auch beim Ansinnen, die Effizienz und Effektivität des Einsatzes öffentlicher Mittel zu prüfen, auf das als „esoterisch“ geltende Wissen der wissenschaftlichen Fachgesellschaften angewiesen sei (Matthies/Simon/Knie 2008).

Weitere häufige Kritikpunkte am Peer Review sind ein Mangel an Transparenz, ein hohes Maß an Subjektivität des Verfahrens, ein geringer Grad an Gutachterübereinstimmung sowie der große Aufwand und die potenzielle Manipulationsanfälligkeit des Verfahrens durch die Auswahl der Expertinnen und Experten (*Bornmann 2004, S. 12f.*). Im Verhältnis zur Wissenschaft erlange der Staat durch das Peer Review nur bedingt eine Kontrollkompetenz, da das Urteil über die wissenschaftliche Leistungskraft letztlich der Wissenschaft selbst überlassen bleibe. Die impliziten Kriterien der Bewertung von Wissenschaft führten somit zu einer „gefühlten Exzellenz“. Der Versuch, wissenschaftspolitische Entscheidungen durch ein Peer Review-Verfahren zu legitimieren, sei also ein strukturelles Dilemma der Wissenschaftspolitik (*Matthies/Simon/Knie 2008*). Weiterhin dürften Peers nicht als „passive Messinstrumente“ missverstanden werden, da diese beispielsweise durch bestimmte theoretische Ansätze voreingenommen oder versucht sein könnten, so zu urteilen, dass Konflikte vermieden werden. Die Begutachteten wiederum könnten geneigt sein, taktisch zu agieren, wenn die Bewertungen in einem Zusammenhang mit der Zuweisung finanzieller Mittel stehen, sodass „(im) Fazit wissenschaftliche Bewertungsprozesse (...) eine starke soziale Komponente (besitzen)“ (*Riese 2007, S. 133*). Andererseits sind Begutachtungen auf der Basis von Peer Reviews „ein nicht mehr wegzudenkendes Instrument zur Reduktion der Entscheidungskomplexität und Verbesserung der Entscheidungsgrundlage für eine leistungsorientierte Mittelverteilung sowohl für angehende Studenten und potenzielle Arbeitgeber als auch für den öffentlichen Träger“ (*Riese 2007, S. 133f.*). Auf diese Weise kann auch der Problematik bei der Nutzung von Indikatoren und Kennzahlen begegnet werden, die darin besteht, „dass ihre Anwendung auf die zu bewertenden Projekte, Forscher und Forschungsgruppen jenseits von Routinefällen selber der Prüfung durch Peers bedarf, um zu sichern, dass sie verlässliche und valide Indikatoren der Leistungen darstellen, die gemessen werden sollen“ (*Neidhardt 2010, S. 290*).

5 Ziele und Anlass des sächsischen Hochschulratings

Einen wesentlichen Anlass für das sächsische Hochschulrating bildeten die erwarteten Auswirkungen der demographischen Entwicklung für die Hochschulentwicklungsplanung in Sachsen, da diese in erheblichem Maße die Studiennachfrage determiniert. In Sachsen wirkt sich der demographische Wandel den zu Projektbeginn vorliegenden Prognosen und Untersuchungen zufolge wesentlich dramatischer aus als in vielen westdeutschen Bundesländern sowie in den Stadtstaaten.¹ Grund hierfür sind zwei drastische demographische Einschnitte nach der politischen Wende im Jahr 1989: der Geburteneinbruch während der Wiedervereinigungsphase (der sogenannte „Geburten-

¹ Es ist an dieser Stelle nicht möglich, die zwischenzeitlich erschienenen Prognosen zur Entwicklung der Studiennachfrage zu berücksichtigen, die von einer deutlich höheren Zahl an Studienanfängerinnen und -anfängern ausgehen, und die damaligen Annahmen entsprechend zu hinterfragen. Die Güte von Prognosen ist als strukturelles Problem von Planungsprozessen insgesamt zu betrachten und stellt somit keinen spezifischen Aspekt des vorgestellten Verfahrens dar.

knick“) und die Abwanderung. Bereits ab 1980, aber vor allem in den Jahren nach der Wende, verzeichnete Sachsen einen erheblichen Geburtenrückgang auf nur noch gut ein Drittel des Ausgangsniveaus. Hinzu kommt ein Wanderungsverlust, der insbesondere in den Jahren 1990 und 1991 sehr hoch war und nach 1998 erneut auftrat. So hat sich zwischen 1990 und 2005 die Zahl der Kinder und Jugendlichen unter 15 Jahren halbiert. Bis 2020 wird der Bestand in dieser Altersgruppe in etwa gleich bleiben. Diese deutlich reduzierten Jahrgänge bilden die Studienanfängerjahrgänge der Jahre ab 2010. Die Hochschulen müssten deshalb mit einem demographisch bedingten Rückgang der Studienanfängerzahlen rechnen, so die Annahme.

Im Ergebnis galten unterschiedliche Prognosen der Entwicklung der Studienanfängerzahlen als mögliche Szenarien, die unter der Annahme einer weitgehenden Fortschreibung des Status quo einen Rückgang der Studienanfängerzahlen von 21.792 Personen im Ausgangsjahr 2003 bis zum Jahr 2015 um 41 bzw. 49 Prozent auf 12.800 bzw. 11.100 Personen für möglich hielten (*Lenz u. a. 2007, S. 72ff.*). Ein so großer Rückgang der Studienanfängerzahlen galt aber als unwahrscheinlich, da Politik und Hochschulen verschiedene Instrumente zur Verfügung stehen, um den zu erwartenden Rückgang abzuschwächen. Dies gilt insbesondere für eine Erhöhung der Studierneigung von jungen Frauen und die gezielte Anwerbung von westdeutschen und ausländischen Studienberechtigten. Als realistisch galt deshalb bis spätestens 2015 ein perspektivischer Rückgang der Studienanfängerzahlen um etwa 25 Prozent gegenüber dem Jahr 2005 und somit von ca. 20.000 auf ca. 15.000 Studienanfängerinnen und -anfänger (*Berthold/Hener/von Stuckradt 2008, S. 7ff.; Lenz u. a. 2007, S. 72ff.; Sächsischer Landtag 2008, S. 2f.*). In der Folge eines Rückgangs der Studienanfängerzahlen um 25 Prozent wären Anpassungen beim Studienangebot und weitergehende Strukturentscheidungen nicht zu vermeiden. Insbesondere für Fächer, bei denen bereits vor der Fortschreibung der Hochschulentwicklungsplanung des Freistaates Sachsen die Studiennachfrage rückläufig und die Auslastung der Lehreinheiten nicht zufriedenstellend waren, stellte sich aus einer landesplanerischen Perspektive die Frage nach einer zukunftsfähigen Entwicklungsstrategie. Analog zur Entwicklung der Studienanfängerzahlen wurde erwartet, dass dieselben Trends zeitversetzt auf die Absolventenzahlen zutreffen würden. Die zu erwartenden Defizite beim Fachkräfteangebot wurden auf 50.000 (*Ragnitz u. a. 2007*) bis 90.000 Personen (*Grundig/Pohl 2007*) im Jahr 2020 geschätzt.

Vor diesem Hintergrund sahen sich die sächsischen Hochschulen – genauso wie die staatliche Hochschulentwicklungsplanung – vor einer doppelten Herausforderung: Angesichts der Gefahr einer abnehmenden Studienanfängerzahl bei gleichbleibend hohem Bedarf an hochqualifizierten Fachkräften galt es einerseits, die überregionale Attraktivität der sächsischen Hochschulen für Studierende zu stärken und zusätzliche Studierendengruppen zu gewinnen. Andererseits sollte gleichzeitig die Erfolgsquote

der bisherigen Studierendengruppen ohne Qualitätseinbuße erhöht werden. Das sächsische Hochschulrating legte deshalb ein besonderes Augenmerk auf die Entwicklung der Studiennachfrage, der Erfolgsquote sowie der Anteilswerte von Studienanfängerinnen und ausländischen Studierenden. Die Begutachtung sollte zeigen, ob die einzelnen Hochschulen auf der Ebene der Studienfächer auf die Herausforderungen eingestellt sind, die sich aus dem demographischen Wandel ergeben. Die Expertinnen und Experten erweiterten ihre Betrachtung um bestehende Kontextbedingungen und mögliche Wechselwirkungen und nahmen eine differenzierte Analyse der Stärken und Schwächen der einzelnen Fächer vor, wobei sie im Rahmen der bestehenden Möglichkeiten auch die jeweiligen Chancen und Risiken berücksichtigten. Eine solche umfassende Betrachtung, die mehrere Bewertungsdimensionen einschließt, sollte eine zu starke Unterwerfung der Hochschulentwicklungsplanung unter wahrgenommene oder tatsächlich bestehende Sachzwänge verhindern und so eine belastbare Ausgangsbasis für eine evidenzbasierte wissenschaftliche Landesplanung bilden.

6 Vorgehensweise

Das sächsische Hochschulrating verbindet eine umfassende indikatorbasierte Analyse von ausgewählten Fächern mit einem qualitativen Bewertungsprozess.² Das Verfahren wurde zwischen Mai 2008 und April 2009 im Auftrag des Sächsischen Staatsministeriums für Wissenschaft und Kunst³ für 21 Studienfächer⁴ an zehn sächsischen Hochschulen⁵ durchgeführt. Insgesamt wurden somit 116 „Fälle“ in die Betrachtung einbezogen. Grundlage der Analyse bildete eine vergleichende Betrachtung von insgesamt 23 Indikatoren aus den vier Bereichen „Studium und Lehre“, „Forschung und wissenschaftlicher Nachwuchs“, „Ausstattung und monetäre Kennzahlen“ sowie „Transfer in andere gesellschaftliche Bereiche“. Die Bewertung der einzelnen Fächer erfolgte im Sinne einer qualitativen Bilanzierung der Stärken und Schwächen nicht durch das

² Ursprünglich war vorgesehen, die Bewertung der verschiedenen Dimension analog zum Forschungsrating des Wissenschaftsrates nach folgender Skala: „Exzellent“, „Sehr gut“, „Gut“, „Befriedigend“ und „Nicht befriedigend“ sowie „Nicht bewertbar“ durchzuführen. Die Expertinnen und Experten hielten allerdings die im Rating vorgesehene Bewertung nach Punkten ohne eine Benennung der Ursachen für nicht empfehlenswert. Das Verfahren wurde daraufhin verändert.

³ Die Trägerschaft wurde durch die Sächsische Akademie der Wissenschaften bzw. das Sächsische Kompetenzzentrum für Bildungs- und Hochschulplanung an der TU Dresden übernommen.

⁴ Hierbei handelte es sich um die folgenden Fächer: Geschichte, Erziehungswissenschaften, Sozialpädagogik/ Sozialwesen, Germanistik, Romanistik, Slawistik, Sport/ Sportwissenschaft, Betriebswirtschaftslehre/ Betriebswirtschaft, Wirtschaftsingenieurwesen, Mathematik, Biologie, Chemie, Physik, Informatik, Maschinenbau/Maschinenwesen, Verfahrenstechnik, Fahrzeugtechnik, Mechatronik, Elektrotechnik/Informationstechnik, Bauingenieurwesen und Architektur.

⁵ Hierbei handelte es sich um alle staatlichen Universitäten und Fachhochschulen, also die Technischen Universitäten in Chemnitz, Dresden und Freiberg, die Universität Leipzig sowie die Fachhochschulen in Dresden, Mittweida, Leipzig, Zittau/Görlitz und Zwickau. Hinzu kam die Evangelische Hochschule Dresden, die auf eigenen Wunsch in die Betrachtung einbezogen wurde.

Ministerium selbst, sondern durch 17 unabhängige Fachexpertinnen und -experten mit großer wissenschaftspolitischer Expertise. Hierbei handelte es sich um Hochschul-lehrerinnen und Hochschullehrer von Universitäten und Fachhochschulen, die nicht an sächsischen Hochschulen tätig waren und über Erfahrungen aus Leitungsfunktionen an Hochschulen sowie Gremien von bundesweiter Relevanz wie beispielsweise dem Wissenschaftsrat, dem Senat der Deutschen Forschungsgemeinschaft oder dem Akkreditierungsrat verfügten. Beim sächsischen Hochschulrating handelt es sich somit nicht um eine formelbasierte Berechnung, wie dies in vielen Rankings der Fall ist, sondern um eine Peer-Review-Bewertung unter Rückgriff auf quantitative und qualitative Daten. Diese Vorgehensweise im Sinne eines Methodenmixes hatte zum Ziel, die Probleme der „unreflektierten Nutzung von Indikatoren“ und „entkontextualisierten Informationen“ bei quantitativen Verfahren bzw. das Risiko eines „subjektiven Urteils-bias“ bei den Anhörungen zu minimieren und durch die Unabhängigkeit der Bewertungen die Akzeptanz der Hochschulen sicherzustellen (Hornbostel 2010, S. 304).

Der Bewertungsprozess umfasste drei Phasen: In der ersten Phase wurden quantitative und qualitative Daten erhoben. Für jeden Indikator und jedes Fach wurden die Daten aller beteiligten Hochschulen ausgewiesen, was die Möglichkeit eines hochschulübergreifenden Vergleichs schuf. Der Erhebungszeitraum von fünf Jahren (2003 bis 2007) ermöglichte die Übersicht über spezifische Entwicklungen innerhalb der Bewertungsdimensionen in diesem Zeitraum. Allen beteiligten Hochschulen wurde der vollständige Datenkatalog übermittelt und die Gelegenheit gegeben, zur vergleichenden indikatorbasierten Betrachtung der Fächer eine Stellungnahme abzugeben, um auf diese Weise Zusammenhänge zwischen den quantitativen Daten herzustellen, bestimmte Ausprägungen zu erklären und qualitative Ergänzungen zu leisten. Es sei an dieser Stelle ausdrücklich darauf hingewiesen, dass sowohl im wissenschaftlichen als auch im hochschulpolitischen Diskurs „kein Konsens über die Operationalisierung der Qualität wissenschaftlicher Leistungen“ besteht, was sich beispielsweise bei der Diskussion von internationalen Rankings und Ratings deutlich gezeigt hat. Auch wurden „bisher kaum theoretische Begründungen für die Auswahl und Gewichtung von Indikatoren vorgebracht“ (Lange 2010, S. 326). Dieses Defizit konnte auch in dem beschriebenen Projekt nicht behoben werden. Die meisten Indikatoren des sächsischen Hochschulratings entsprechen vielmehr den „üblichen“ Indikatoren der leistungsbezogenen Bewertungsverfahren (Lüthje 2010, S. 269) und folgen somit weniger einer expliziten Begründung ihrer Zweckmäßigkeit als vielmehr einer für den hochschulpolitischen Diskursraum begründet vermuteten, also letztlich impliziten Normativität.

Basierend auf der indikatorgestützten Bewertung und den Stellungnahmen der Hochschulen bzw. Fachbereiche wurde in einem zweiten Schritt eine Peer-Review-Bewertung durchgeführt, in deren Rahmen Anhörungen mit Vertreterinnen und Vertretern

aller Hochschulen bzw. Fachbereiche durchgeführt wurden. Dies ermöglichte eine (qualitative) Validierung der indikatorgestützten Bewertung, sicherte die Transparenz des Verfahrens und bot den Beteiligten gewisse Mitwirkungsmöglichkeiten. Der Bewertungsprozess erfolgte fächergruppenbezogen in drei getrennten Verfahren („Cluster“), um die Belastung für die externen Expertinnen und Experten zu begrenzen und die Akzeptanz der Bewertung aufgrund der fachlichen Nähe der Externen zu erhöhen.

In der dritten Phase erfolgte in der Verantwortung der Expertinnen und Experten die Erstellung eines Stärken- und Schwächen-Profiles des jeweiligen Fachs im Kontext der anderen Angebote an den sächsischen Hochschulen. Hierzu wurde zunächst eine clusterübergreifende Redaktionsgruppe gebildet, deren Berichtsentwurf im Umlaufverfahren mit allen externen Sachverständigen abgestimmt wurde. Über eine Matrixstruktur konnte sowohl ein Überblick über die Ratingfächer nach Standorten als auch über die beteiligten Hochschulen und deren Angebot in den drei Clustern gegeben werden. Auf diese Weise ließ sich eine Vergleichbarkeit zwischen den Clustern, innerhalb der Cluster und zwischen den Standorten herstellen. Auch hier wurde den Hochschulen die Möglichkeit zur sachlichen Korrektur bzw. Ergänzung durch Stellungnahmen gegeben. Neben den Teilberichten zu den Ratingfächern und den Standorten enthält der Bericht auch ein Einleitungskapitel mit den Motiven und einer Beschreibung der Herangehensweise sowie generelle Aussagen zu den Fächern und Standorten.

7 Indikatoren und Merkmale

Um das Ziel einer umfassenden Bewertung der Stärken und Schwächen der einbezogenen Fächer zu erreichen, wurden vier Bewertungsdimensionen festgelegt, die jeweils durch unterschiedliche Indikatoren unterlegt wurden. Einen Überblick bietet die nachstehende Tabelle.

Tabelle: Bewertungsdimensionen und Indikatoren des sächsischen Hochschulratings

Dimensionen	Indikatoren
Studium und Lehre	angebotene Studiengänge Ergebnisse von Akkreditierungsverfahren Zulassungsbeschränkungen Anzahl der Studienanfängerinnen und -anfänger Anzahl der Studierenden Anteil der Studierenden in der Regelstudienzeit Anzahl der Absolvierenden und Absolventen durchschnittliche Fachstudiendauer Anzahl der Absolvierenden und Absolventen innerhalb der Regelstudienzeit
Forschung und wissenschaftlicher Nachwuchs	Summe der Drittmittel Anzahl der Promotionen Beteiligung an Graduiertenkollegs Beteiligung an strukturierten Forschungsprojekten Forschungspreise
Transfer in andere gesellschaftliche Bereiche	Tätigkeit als Fachgutachterinnen und -gutachter erfolgreiche Patentanmeldungen weiterbildende Studienangebote Weiterbildungsteilnehmerinnen und -teilnehmer Gasthörerinnen und Gasthörer
Ausstattung und monetäre Kennzahlen	Anzahl der Professorinnen und Professoren Anzahl der wissenschaftlichen Beschäftigten (jeweils als Fallzahlen und als Vollzeitäquivalente) Auslastung der zugeordneten Lehrereinheit Summe der laufenden Ausgaben

Eine Besonderheit des vorgestellten Verfahrens ist die Einbeziehung der qualitativen Indikatoren Akkreditierungsergebnisse und Zulassungsbeschränkungen. Dies hat folgenden Hintergrund: Da die Hochschulen in Sachsen weder zur Einführung von Bachelor- und Masterstudiengängen noch zur Akkreditierung ihrer Studiengänge verpflichtet sind, wurde dem Indikator Akkreditierungsergebnisse zugeschrieben, einen Hinweis auf die Bereitschaft der sächsischen Hochschulen zu geben, sich an den Vorgaben des Bologna-Prozesses zu orientieren und diese umzusetzen. Von den Angaben zum Vorhandensein von Zulassungsbeschränkungen wurde in Verbindung mit der Auslastung ein Hinweis auf Angebot und Nachfrage der Studienfächer sowie auf die Zulassungspolitik der Hochschulen erwartet.

Um den Aufwand der Datenerhebung für die Hochschulen so gering wie möglich zu halten und gleichzeitig die Vergleichbarkeit der Daten zu sichern, wurden im sächsischen Hochschulrating soweit wie möglich die Definitionen der amtlichen Statistik übernommen. Teilweise wurden Modifikationen vorgenommen, um die Passfähigkeit zwischen den verschiedenen Systematiken der Hochschulstatistiken zu verbessern. Weitere notwendige Anpassungen der Datendefinitionen resultierten aus inhaltlich fehlerhaften Zuordnungen, die teilweise auf Abstimmungsproblemen zwischen den Hochschulen

und dem Statistischen Landesamt Sachsen beruhten, in erster Linie aber durch inhaltlich fehlerhafte Zuordnungen der Hochschulen selbst entstanden sind. Teilweise wurden seitens der Hochschulen auch die ermittelten Daten in Frage gestellt. Die anschließenden Überprüfungen der Fehlermeldungen der Hochschulen führten allerdings in jedem Fall zu dem Ergebnis, dass die unterschiedlichen Daten, die bei den Hochschulen und beim Statistischen Landesamt vorlagen, aus der Anwendung unterschiedlicher Definitionen durch die Hochschulen beziehungsweise fehlerhaften Zuordnungen resultierten. Beispiele hierfür sind die unterschiedliche Behandlung beurlaubter Studierender oder von Studierenden, die nach dem Erhebungsstichtag im laufenden Semester exmatrikuliert wurden. Diese Korrekturen der Daten, die das Statistische Landesamt vornimmt, wurden nicht von allen Hochschulen übernommen. Als Konsequenz aus den Schwierigkeiten bei der Datenerhebung wurde nach der Durchführung des sächsischen Hochschulratings zwischen dem Sächsischen Staatsministerium für Wissenschaft und Kunst, dem Statistischen Landesamt Sachsen und den Hochschulen eine umfassende Überprüfung der fachlichen Zuordnungen vereinbart.

Trotz aller strukturellen und operativen Schwächen der Daten der amtlichen Statistik bietet deren Nutzung mehrere entscheidende Vorteile. So sind die Definitionen bundesweit einheitlich festgelegt, dokumentiert und leicht zugänglich (*Statistisches Bundesamt 2007a, 2007b, 2007c*). Weiterhin werden die gemeldeten Daten von fachlich qualifiziertem Personal geprüft, gegebenenfalls korrigiert und aufbereitet. Die Datenqualität der amtlichen Statistik kann somit insgesamt als hoch gelten. Anhand der genannten Beispiele sollte hinreichend deutlich geworden sein, dass bei vergleichenden Untersuchungen die Nutzung der Daten der amtlichen Statistik trotz aller Einschränkungen einer Nutzung von Daten, die durch die Hochschulen selbst ermittelt wurden, grundsätzlich vorzuziehen ist.

8 Empfehlungen der Expertinnen und Experten

Zwischen dem sächsischen Wissenschaftsministerium und den Hochschulen wurde für die fachspezifischen Inhalte des Expertenberichts Vertraulichkeit vereinbart. Es ist deshalb an dieser Stelle nicht möglich, nähere Ausführungen zu den Bewertungen der einzelnen Fächer oder zu übergreifenden Strukturempfehlungen zu machen.⁶ Die Darstellung muss sich also auf vergleichsweise allgemeine Aussagen beschränken. So haben die Expertinnen und Experten zunächst klargestellt, dass sie ihre Aufgabe nicht darin sehen, kurzfristige Kürzungspotenziale zu identifizieren. Vielmehr sei die Zielsetzung, inhaltliche und strukturelle Vorschläge zu unterbreiten, die vor dem Hintergrund der Herausforderungen des demographischen Wandels zu einer Stärkung der sächsi-

⁶ Die Empfehlungen der Expertinnen und Experten zu Anpassungen bei der Struktur des Studienangebots an den sächsischen Hochschulen wurden im Wesentlichen in den sächsischen Hochschulentwicklungsplan übernommen und können dort nachgelesen werden (*SMWK 2011*).

schen Hochschul- und Forschungslandschaft beitragen könnten. Hierzu gehören nach Einschätzung der Expertinnen und Experten vor allem Aktivitäten in den Bereichen der Cluster- und Profilbildung. Dies beinhaltet insbesondere eine stärkere Kooperation der Hochschulen sowie eine stärkere Konzentration des Studienangebots in bestimmten Fächern. Vor allem Hochschulen in unmittelbarer räumlicher Nähe sollten ihr Studienangebot stärker aufeinander abstimmen und beispielsweise die wechselseitige Anerkennung von Modulen gewährleisten. Dies gilt nach Ansicht der Expertinnen und Experten auch für die Zusammenarbeit von Universitäten und Fachhochschulen, die überdies durch die Einrichtung von gemeinsamen Graduiertenkollegs unterstützt werden sollte. Außerdem stehe es außer Frage, dass der demographisch bedingte Rückgang der Studienanfängerzahlen hinsichtlich des bestehenden Fachkräftebedarfs letztlich nur kompensiert werden könne, wenn einerseits die Erfolgsquote der Studierenden deutlich steige, was insbesondere eine Aufgabe der Natur- und Ingenieurwissenschaften sei, und andererseits zusätzliche Zielgruppen zur Aufnahme eines Studiums motiviert werden könnten. Hierzu gehören den Expertinnen und Experten zufolge insbesondere westdeutsche und ausländische Studienberechtigte, beruflich Qualifizierte sowie weibliche Studienberechtigte, die in Sachsen eine unterdurchschnittliche Studierneigung aufwiesen. Für einige Fächer, die bereits in der Vergangenheit vergleichsweise gering ausgelastet waren, regten die Expertinnen und Experten aber auch die Konzentration bzw. Zusammenlegung von Studienangeboten an, was unter anderem durch die Schließung von Außenstandorten erfolgen solle. Denn die ausreichende Zahl an Studierenden und wissenschaftlichem Personal stelle eine notwendige Voraussetzung für die dauerhafte Zukunftsfähigkeit eines Fachs dar. Möglichkeiten zur Konzentration von Studienangeboten wurden insbesondere in den Fächern Bauingenieurwesen und Architektur, Informatik, Betriebswirtschaftslehre und Sozialpädagogik/Sozialwesen sowie teilweise auch in den Fächern Mathematik, Maschinenbau und Elektrotechnik gesehen (*Expertengruppe Sächsisches Hochschulrating 2009, S. 19–23, S. 153–157*).

9 Ertrag für die Hochschulentwicklungsplanung

Trotz des zwischenzeitlich erfolgten Regierungswechsels von einer Koalition von CDU und SPD zu einer Koalition von CDU und FDP und dem damit verbundenen Wechsel im Amt der Wissenschaftsministerin zeigt der am 17. Dezember 2011 von der sächsischen Staatsregierung beschlossene Hochschulentwicklungsplan (SMWK 2011) hinsichtlich der vorgeschlagenen Strukturentscheidungen Akzentuierungen, die sehr weitgehend mit den Ergebnissen und Empfehlungen des sächsischen Hochschulratings übereinstimmen. Die Inhalte des Hochschulentwicklungsplans wurden bereits im Laufe des Jahres 2011 in der Öffentlichkeit diskutiert. So hatte die sächsische Wissenschaftsministerin Sabine von Schorlemer im Mai 2011 eine stärkere Kooperation zwischen den Hochschulen sowie den Abbau von „unnötigen Doppelangeboten“ als zentrale Ziele der Hochschulentwicklungsplanung benannt. Diese Vorgehensweise folgte in

erster Linie hochschulpolitischen Erwägungen. Struktur- oder regionalpolitischen Aspekten kam eine allenfalls untergeordnete Bedeutung zu. Beispielsweise sollten die Fächer Architektur und Bauingenieurwesen künftig in Dresden und Leipzig konzentriert werden, was den Wegfall der entsprechenden Angebote an den Standorten in Zittau und Reichenbach nach sich ziehen würde. Wegfallen sollen ebenfalls die Außenstellen der Fachhochschulen Mittweida in Roßwein und Zwickau in Reichenbach. Hinzu kommt, dass das Internationale Hochschulinstitut (IHI) in Zittau seinen bisherigen Status als „Mini-Universität“ verlieren und organisatorisch in die Bergakademie Freiberg überführt werden soll. In den Fächern Informatik und Wirtschaftswissenschaften sollen weiterhin insbesondere die Universitäten und Fachhochschulen in Leipzig und Dresden gegenseitig verstärkt Lehrmodule anerkennen und kooperative Promotionsverfahren häufiger nutzen (*Pfeil 2011; Schön 2011*). Insbesondere die geplante Schließung von Studiengängen bzw. Außenstandorten zog erwartungsgemäß heftige Proteste in den betroffenen Städten nach sich. So protestierten in Zittau, Roßwein und Reichenbach Studierende, aber auch Lokalpolitikerinnen und -politiker sowie die Landtags- und teilweise auch die Bundestagsabgeordneten aus der Region. Hierbei sind für Sachsen ungewöhnliche parteiübergreifende Bündnisse zu beobachten – beispielsweise von einzelnen Abgeordneten der CDU und der Linkspartei. Heftigen Widerspruch erntete Ministerin von Schorlemer auch bei den Landtagsfraktionen von SPD und Grünen (*Lemke 2011; Lotze 2011; Zenker 2011*). Vor diesem Hintergrund ist es auf den ersten Blick überraschend, dass die Reaktionen der betroffenen Hochschulleitungen auf die Planungen des Wissenschaftsministeriums tendenziell eher zustimmend ausfielen. So erklärte beispielsweise Michael Hösel, Prorektor der Hochschule Mittweida, dass Roßwein schon immer eine „Notlösung“ gewesen sei und allein schon die vorhandene Bausubstanz einen Weiterbetrieb des Standortes kaum zulasse (*Lemke 2011*). Der Rektor der Hochschule Zittau/Görlitz, Friedrich Albrecht, erklärte, dass er vor dem Hintergrund der demographischen Entwicklung die Entscheidung zur Schließung der Zittauer Bau-Fakultät nachvollziehen könne, auch wenn er diese aus Hochschulsicht bedaure (*Zenker 2011*). Der Rektor der Westsächsischen Hochschule Zwickau, Günter Krautheim, ging noch einen Schritt weiter und bekannte sich öffentlich zum „Konzept für ein geordnetes Ende“ des Standortes Reichenbach. Der Trend rückläufiger Studierendenzahlen habe unabhängig von den jüngst veröffentlichten Schließungsplänen des Ministeriums ein Nachdenken über die Zukunft des Außenstandortes notwendig gemacht. Gleichzeitig sprach sich Krautheim auch gegen eine Überführung des Faches Architektur an die Berufsakademie in Plauen aus, da ein solches Angebot kaum auf die notwendige Anerkennung stoßen würde (*Möckel 2011*). Die übereinstimmende Zustimmung der Hochschulleitungen könnte jedoch auch daraus resultieren, dass ihnen die möglichen Strukturentscheidungen bereits seit mehr als zwei Jahren bekannt waren und sie diese als Ergebnis eines transparenten und partizipativen Bewertungsprozesses auf der Basis des sächsischen Hochschulratings insgesamt anerkennen.

10 Erfahrungen mit dem sächsischen Hochschulrating

Das sächsische Hochschulrating hat gezeigt, dass die Zusammenschau verschiedener quantitativer und qualitativer Indikatoren wichtige Hinweise für eine Bewertung der Stärken und Schwächen eines Faches bieten kann, die eine hinreichende Grundlage für die Zwecke der staatlichen Hochschulentwicklungsplanung bilden können. Hierbei hat es sich als essentiell erwiesen, dass die erhobenen Daten durch sachkundige Personen interpretiert werden und erst auf dieser Basis eine Bewertung der Leistungen eines Fachs erfolgt. Wichtig hierbei ist auch die Einbeziehung der Vertreterinnen und Vertreter der zu bewertenden Fächer sowie der Verantwortlichen in den Hochschulverwaltungen sowohl bei der Datenerhebung als auch bei der Formulierung der jeweiligen Situationsbeschreibungen. Erst dieser Zwischenschritt der gemeinsamen Dateninterpretation ermöglicht es, Datenfehler im Detail zu ermitteln, und besondere Umstände wie einen generationsbedingten Wechsel beim wissenschaftlichen Personal oder den Umgang mit zusätzlichen Aufgaben wie der Umsetzung des Hochschulpaktes angemessen zu berücksichtigen. Die Möglichkeit zur Formulierung begründeter Einwände trägt auch zur Akzeptanz des Verfahrens bei, selbst wenn diese sich bei genauerer Prüfung nicht selten als unbegründet erwiesen.

Ein weiteres Ergebnis des sächsischen Hochschulratings war, dass die „üblichen“ Indikatoren nicht immer auch die „richtigen“ Indikatoren sein müssen. So haben die Expertinnen und Experten das erfolgreiche Durchlaufen eines Akkreditierungsverfahrens in manchen Fällen als aussagekräftiger für die Bewertung der Leistung eines Fachs im Bereich Studium und Lehre eingeschätzt als beispielsweise die Angaben zur durchschnittlichen Fachstudiendauer. Ähnlich waren die Einschätzungen im Bereich der Forschung. Hier wurden die Aktivitäten des wissenschaftlichen Personals in Graduiertenkollegs und die Beteiligung an strukturierten Forschungsprojekten, durchaus auch im Verbund mehrerer Projektpartner, als aussagekräftiger für die Leistungsbewertung betrachtet als die Anzahl der Promotionen oder die schlichte Höhe der eingeworbenen Drittmittel. Dieser Eindruck ist sicherlich nicht umstandslos verallgemeinerungsfähig, lädt aber dennoch dazu ein, bereits bei der Konzeption von Verfahren der externen Leistungsbewertung bzw. der Qualitätssicherung intensiv zu überlegen, welche Indikatoren tatsächlich dafür geeignet sind, als Qualitätsmaßstab gelten zu können und bei welchen Indikatoren auf eine gesonderte Erhebung eher verzichtet werden kann. Zu den Indikatoren, die sich im Rahmen des Projekts nicht bewährt haben und somit für die formulierte Aufgabenstellung als entbehrlich gelten können, gehören unter anderem die fachgutachterliche Tätigkeit des wissenschaftlichen Personals, die von diesem erhaltenen Forschungspreise sowie das Vorhandensein von Zulassungsbeschränkungen. Eingeschränkt gilt dies auch für die Anzahl der Patente, die für sich genommen noch keine belastbaren Aussagen über die Leistungen eines Fachs im Bereich des Wissenstransfers geben kann. Für den Bereich der Weiterbil-

dungsangebote muss weiterhin konstatiert werden, dass die Studierendenstatistik der Statistischen Ämter hierfür bislang keine geeignete Datengrundlage bietet, da hier nur Studierende in weiterbildenden Studiengängen ermittelt werden können, die Teilnehmerinnen und Teilnehmer von Zertifikatskursen oder einzelnen Modulen also nicht zwingend berücksichtigt werden. Hinzu kommt, dass die Einführung von weiterbildenden Masterstudiengängen zumindest bislang von der Systematik der amtlichen Statistik nicht nachvollzogen worden ist.

Bei der Auswahl der Expertinnen und Experten wurde deren Sachkunde im Umgang mit konkreten Fragestellungen der Hochschulentwicklungsplanung als entscheidendes Kriterium festgelegt. Insbesondere die Erfahrungen in hochschulpolitischen Gremien und als Mitglied einer Hochschulleitung wurden als wichtige Faktoren identifiziert. Bei der Zusammensetzung der Expertengruppe wurde mit dem sächsischen Hochschulrating insofern ein eigener Weg beschritten, als die Sachkunde der Expertinnen und Experten gegenüber ihrer hochschul- und wissenschaftspolitischen Reputation oder der Pluralität ihrer Zusammensetzung von Beginn an höher gewichtet wurde. Diese Vorgehensweise hat sich insgesamt bewährt, da die Bewertungen der Expertinnen und Experten aufgrund ihrer hohen fachlichen Nähe und der deutlich stärker praxisorientierten als programmatischen Perspektive bei den beteiligten Vertreterinnen und Vertretern der Hochschulen und Fachbereiche eine vergleichsweise hohe Akzeptanz erzielten.

Als entscheidend für die Akzeptanz des Verfahrens und dessen Ergebnisse hat sich weiterhin das hohe Maß an interner Transparenz erwiesen. So verfügten alle Beteiligten in jeder Projektphase über sehr viele Informationen – sowohl hinsichtlich ihres eigenen Fachs bzw. ihrer eigenen Hochschule als auch hinsichtlich der anderen in die Bewertung einbezogenen Standorte. Bereits im Vorfeld wurden die Hochschulen über die Zielsetzung, die geplante Verfahrensweise und die vorgesehenen Expertinnen und Experten informiert. Alle Hochschulen verfügten am Ende der Phase der Datensammlung über den vollständigen Datenkatalog, und ihnen wurde die Möglichkeit gegeben, hierzu eine Stellungnahme abzugeben. Im Anschluss an die Anhörungen durch die Expertengruppe erhielten die Hochschulen deren Entwürfe zu den festgestellten fachlichen Profilen der berücksichtigten Fächer ihrer Hochschulen und ebenfalls die Gelegenheit zur Stellungnahme. Der clusterübergreifend abgestimmte Expertenbericht wurde nach der Übergabe an die Staatsministerin für Wissenschaft und Kunst mit den Rektoren der sächsischen Hochschulen diskutiert und erörtert. Eine offizielle Information der Öffentlichkeit ist aufgrund der zwischen dem sächsischen Wissenschaftsministerium und den Rektoren vereinbarten Vertraulichkeit bisher nicht erfolgt.

11 Schlussfolgerungen aus dem Modellprojekt

Insgesamt konnten die mit dem sächsischen Hochschulrating verbundenen Ziele weitgehend erreicht werden. So war es durch die Integration der quantitativen Datenerhebung in die Arbeit einer aus Peers zusammengesetzten Expertenkommission möglich, trotz der geringen Analysetiefe eine vergleichsweise umfassende Leistungsbewertung der berücksichtigten Fächer aus der spezifischen Perspektive der Landeshochschulplanung vorzunehmen. Weiterhin spielten die unmittelbaren Interessen von Politik und Verwaltung bei dem gewählten Bewertungsverfahren eine untergeordnete Rolle. Dies ist nicht nur ein wichtiger Beitrag für eine aus fachwissenschaftlicher Perspektive angemessene Berücksichtigung der spezifischen Besonderheiten der einzelnen Fächer, sondern auch eine wichtige Ergänzung der vielfach durch Sachzwänge bestimmten Perspektive der sächsischen Staatsregierung. Bewährt hat sich auch die Auswahl von Sachverständigen, die über eine hohe Expertise in den Bereichen Qualitätssicherung, Hochschulsteuerung und Hochschulentwicklung verfügen. Dies hat erheblich zur Akzeptanz des Verfahrens und seiner Ergebnisse beigetragen. Das sächsische Hochschulrating hat ebenfalls gezeigt, dass die von Politik und Verwaltung teilweise erwarteten Möglichkeiten einer ausschließlich datenbasierten Leistungsbewertung oder einer rein datenbasierten Hochschulentwicklungsplanung in der Praxis als realitätsfern gelten müssen, da zum einen eine so weitgehende Reduzierung der Komplexität kaum zu erreichen ist und zum anderen die verfügbaren Indikatoren aus systematischen und strukturellen Gründen häufig nur eine teilweise Abbildung der bestehenden Realität leisten können.

Durch die Kombination einer umfangreichen Datenanalyse mit einem qualifizierten Peer Review ist es möglich, die staatliche Hochschulentwicklungsplanung auf eine evidenzbasierte Grundlage zu stellen. Dies ist allerdings mit einem erheblichen personellen, sächlichen und finanziellen Aufwand verbunden. Aus diesem Grund konnte mit den einbezogenen 21 Studienfächern auch nur ein kleiner Teil des gesamten Studienangebots der sächsischen Hochschulen berücksichtigt werden. Eine Ausweitung des vorgestellten Bewertungsverfahrens in einer modifizierten Form wäre fachlich sicherlich wünschenswert, aber aufgrund der bestehenden Ressourcenausstattung der Wissenschaftsministerien der Länder wohl auch perspektivisch kaum leistbar. Realistischer, wenn auch sicherlich nicht einfach zu erreichen, dürfte es sein, die strategische Dimension der Hochschulentwicklungsplanung als Querschnittsaufgabe in den Wissenschaftsministerien der Länder zu verankern und hierbei kontinuierlich Datenauswertungen im Sinne eines Hochschulmonitorings zu berücksichtigen. Hierzu wäre es allerdings erforderlich, zunehmend wirtschafts- und sozialwissenschaftlich qualifiziertes Personal für die Berufstätigkeit in einem Landesministerium zu gewinnen. Flankierend hierzu könnte die dauerhafte Einrichtung eines Expertengremiums analog zu den auf Bundesebene bestehenden Sachverständigenräten erfolgen, um die

Ministerien bei der Entwicklung, vor allem der kontinuierlichen Weiterentwicklung einer Entwicklungsstrategie zu unterstützen.

Literatur

Berthold, Christian / Hener, Yorck / von Stuckrad, Thimo (2008): Demografische Entwicklung und Hochschulen – Pilotprojekt Sachsen. Bestandsaufnahme und politische Empfehlungen. CHE-Arbeitspapier Nr. 104, März 2008, Gütersloh.

Bornmann, Lutz (2004): Stiftungspropheten in der Wissenschaft. Zuverlässigkeit, Fairness und Erfolg des Peer-Review, Münster.

Expertengruppe Sächsisches Hochschulrating (2009): Herausforderungen an eine zukunftsfähige Entwicklung der sächsischen Hochschullandschaft bis 2020. Sächsisches Hochschulrating 2008, Bericht im Auftrag des Sächsischen Staatsministeriums für Wissenschaft und Kunst, unveröffentlichter Evaluationsbericht, Dresden.

Grundig, Beate / Pohl, Carsten (2007): Demografischer Wandel in Ostdeutschland: Fluch oder Segen für den Arbeitsmarkt?, Ifo Dresden berichtet, Nr. 3/2007, Dresden.

Heinze, Rolf G. (2009): Rückkehr des Staates? Politische Handlungsmöglichkeiten in unsicheren Zeiten, Wiesbaden.

Hepp, Gerd F. (2011): Bildungspolitik in Deutschland. Eine Einführung, Wiesbaden.

Hornbostel, Stefan (2010): (Forschungs-) Evaluation, in: Simon, Dagmar / Knie, Andreas / Hornbostel, Stefan (Hrsg.): Handbuch Wissenschaftspolitik Wiesbaden, S. 293–309.

Lange, Rainer (2010): Benchmarking, Rankings und Ratings, in: Simon, Dagmar / Knie, Andreas / Hornbostel, Stefan (Hrsg.): Handbuch Wissenschaftspolitik, Wiesbaden, S. 323–334.

Lemke, Uwe (2011): Hochschulpläne rütteln Politiker wach. Hochschulstandort Mittweida soll gestärkt werden, in: Chemnitzer Freie Presse, Lokalausgabe Mittweida vom 8. Juni 2011, S. 9.

Lenz, Karl u. a. (2007): Hochschulen im demografischen Wandel. Die Lage in Sachsen. Bericht im Auftrag des Sächsischen Staatsministeriums für Wissenschaft und Kunst, Dresden.

Lotze, Maria (2011): Was passiert mit der Hochschule in Roßwein?, in: Sächsische Zeitung, Lokalausgabe Döbeln, vom 14. Mai 2011, S. 21.

Lüthje, Jürgen (2010): Aktivierendes Wissenschaftsmanagement, in: Simon, Dagmar / Knie, Andreas / Hornbostel, Stefan (Hrsg.): Handbuch Wissenschaftspolitik, Wiesbaden, S. 262–279.

Matthies, Hildegard / Simon, Dagmar / Knie, Andreas (2008): „Gefühlte“ Exzellenz – Implizite Kriterien der Bewertung von Wissenschaft als Dilemma der Wissenschaftspolitik, in: Matthies, Hildegard / Simon, Dagmar (Hrsg.): Wissenschaft unter Beobach-

tion. Effekte und Defekte von Evaluationen, Leviathan Sonderheft Nr. 24, Wiesbaden, S. 331–343.

Möckel, Gerd (2011): Hochschule: Pläne für geordnetes Aus. Bildungsinstitut hat Vorstellungen zur Abschaffung des Standorts Reichenbach, in: Chemnitzer Freie Presse, Lokalausgabe Reichenbach, vom 13. Juli 2011, S. 9.

Müller, Ulrich u. a. (2010): „Was macht eigentlich ... das Ministerium?“ Überlegungen zu Wissenschaftsministerien von morgen, CHE-Arbeitspapier Nr. 132, Gütersloh.

Neidhardt, Friedhelm (2010): Selbststeuerung der Wissenschaft: Peer Review, in: Simon, Dagmar / Knie, Andreas / Hornbostel, Stefan (Hrsg.): Handbuch Wissenschaftspolitik, Wiesbaden, S. 280–292.

Pasternack, Peer (2003): Hochschulverträge und Zielvereinbarungen. Grundlagen, Modellreferenzen und Berliner Erfahrungen, in: Die Hochschule, Nr. 1/2003, S. 136–159.

Pfeil, Ingolf (2011): Entwicklungsplan: Vorgaben für HTW und TU Dresden / Reichenbach und Roßwein vor der Schließung, in: Dresdner Neueste Nachrichten vom 12. Mai 2011, S. 1.

Ragnitz, Joachim u. a. (2007): Die demografische Entwicklung in Ostdeutschland. Gutachten im Auftrag des Bundesministeriums für Wirtschaft und Technologie, Dresden.

Riese, Karina (2007): Kriterien zur Ressourcensteuerung an Hochschulen, Wiesbaden.

Sächsisches Staatsministerium für Wissenschaft und Kunst (SMWK, 2011). Der Sächsische Hochschulentwicklungsplan bis 2020. Leitlinien und Instrumente für eine zukunftsfähige Entwicklung der sächsischen Hochschullandschaft, Dresden.

Sächsischer Landtag (2008). Antwort der Sächsischen Staatsregierung auf die Große Anfrage der Fraktion Bündnis 90/ Die Grünen im Sächsischen Landtag „Hochschulen in der demografischen Entwicklung“, Drs. 4/10319/2 vom 17.03.2008, Dresden.

Salais, Robert (2008): Evaluation und Politik: Auf der Suche nach guten Indikatoren für die Forschung, in: Matthies, Hildegard / Simon, Dagmar: Wissenschaft unter Beobachtung. Effekte und Defekte von Evaluationen, Leviathan Sonderheft Nr. 24, Wiesbaden, S. 193–212.

Schön, Stephan (2011): Der Freistaat gibt zwei Hochschulstandorte auf, in: Sächsische Zeitung, Lokalausgabe Dresden, vom 12. Mai 2011, S. 8.

Siefken, Sven T. (2007): Expertenkommissionen im politischen Prozess. Eine Bilanz zur rot-grünen Bundesregierung 1998–2005, Wiesbaden.

Statistisches Bundesamt (2007a): Studierende an Hochschulen, Wintersemester 2006/2007, Fachserie 11, Reihe 4.1, Wiesbaden.

Statistisches Bundesamt (2007b): Personal an Hochschulen 2006, Fachserie 11, Reihe 4.4, Wiesbaden.

Statistisches Bundesamt (2007c): Nichtmonetäre hochschulstatistische Kennzahlen 1980–2006, Fachserie 11, Reihe 4.3.1, Wiesbaden.

Wolter, Andrä (2007): From the Academic Republic to the Managerial University – Towards a New Model of Governance in German Higher Education, in: *University Studies* 35, S. 111–132.

Zechlin, Lothar (2007): Strategische Hochschulentwicklung. Überlegungen zu einer Typologie, in: *Die Hochschule*, Nr. 1/2007, S. 115–131.

Zenker, Thomas (2011): Hochschule soll Bau-Fakultät schließen, in: *Sächsische Zeitung*, Lokalausgabe Zittau, vom 24. Mai 2011, S. 13.

Ziegele, Frank (2002): Reformansätze und Perspektiven der Hochschulsteuerung in Deutschland, in: *Beiträge zur Hochschulforschung*, 24. Jahrgang, Heft 3, S. 106–121.

Anschriften der Verfasser:

Ulf Banscherus
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
E-Mail: ulf.banscherus@hu-berlin.de

Mirjam Reiß
Technische Universität Dresden
Sächsisches Kompetenzzentrum für Bildungs- und Hochschulplanung
01062 Dresden
E-Mail: mirjam.reiss@kfbh.de

Ulf Banscherus ist wissenschaftlicher Mitarbeiter in der Abteilung Hochschulforschung des Instituts für Erziehungswissenschaften der Humboldt-Universität zu Berlin. Er war zwischen November 2007 und August 2009 als Projektmitarbeiter im Bereich Hochschulentwicklungsplanung für das Sächsische Staatsministerium für Wissenschaft und Kunst tätig.

Mirjam Reiß ist wissenschaftliche Mitarbeiterin am Sächsischen Kompetenzzentrum für Bildungs- und Hochschulplanung der Technischen Universität Dresden. Sie war zwischen Juni 2008 und Februar 2010 als Projektmitarbeiterin im Bereich Hochschulentwicklungsplanung für das Sächsische Staatsministerium für Wissenschaft und Kunst tätig.

Dill, David D.; Beerkens, Maarja (Hrsg.): Public Policy for Academic Quality. Analyses of Innovative Policy Instruments. Higher Education Dynamics Bd. 30, ISBN 978-90-481-3753-4, 345 Seiten

Der Hochschulbereich hat in den letzten zehn bis 15 Jahren einen tiefgreifenden Wandlungsprozess durchlaufen. Im Zuge dieser Neuordnungen kommt der Qualitätssicherung und Qualitätsentwicklung an Hochschulen eine zentrale Aufgabe zu. Dieser Sammelband präsentiert 14 unterschiedliche Ansätze externer Qualitätssicherung im Hochschulbereich aus acht Ländern (UK, USA, Australien, Hong Kong, Deutschland, Dänemark, Spanien, Brasilien) und fasst damit die Ergebnisse des Forschungsprojekts "Public Policy for Academic Quality Research Program (PPAQ)" zusammen. Die Autoren sind anerkannte Experten in diesem Politik- und Forschungsfeld, unter ihnen neben den Herausgebern Peter T. Ewell, Elaine El-Khawas, Bjørn Stensaker und William F. Massy. Sie stellen die Ansätze und Ziele der nationalen Qualitätsmanagementsysteme vor und gehen auf Probleme bei der Implementierung sowie Wirkungen ein. In einem Abschlusskapitel fassen die Herausgeber die Erfahrungen zusammen und arbeiten die notwendigen Rahmenbedingungen für die Sicherung akademischer Standards im Universitätsbereich heraus.

Hesse, Joachim Jens: Die Internationalisierung der Wissenschaftspolitik: Nationale Wissenschaftssysteme im Vergleich. Berlin: Duncker & Humblot, 2011, ISBN 978-3-428-13672-8, 571 Seiten

Die Internationalisierung der Wissenschafts- und Technologiepolitik bestimmt zunehmend die politische Agenda in Deutschland wie auch in anderen Industrie- und Schwellenländern. Vor dem Hintergrund der „Strategie der Bundesregierung zur Internationalisierung von Wissenschaft und Forschung“ hat das Internationale Institut für Staats- und Europawissenschaften (ISE) im Auftrag des Bundesministeriums für Bildung und Forschung (BMBF) eine Untersuchung zur Internationalisierung der Wissenschaftspolitik durchgeführt. Im Rahmen eines Vergleichs nationaler Wissenschaftssysteme wird der Frage nachgegangen, inwieweit die notwendigen Voraussetzungen für eine verstärkte Internationalisierung der Wissenschaftspolitik auch in anderen Staaten gegeben sind. Die Studie beinhaltet sieben Länderberichte (Finnland, Großbritannien, USA, Japan, Singapur, Volksrepublik China und Indien), in deren Mittelpunkt die Rahmenbedingungen, Institutionen und Akteure der nationalen Wissenschaftspolitiken sowie ihre Ansätze, Ziele und Ergebnisse stehen. Die empirischen Erkenntnisse der Länderstudien stützen sich neben einer reichhaltigen Materialsammlung auf eine Vielzahl von Interviews mit Vertretern der politischen Führung sowie der wissenschaftlichen und wirtschaftlichen Eliten der Vergleichsländer. Die zum Schluss formu-

lierten Leitsätze zeigen Handlungsbedarfe und Empfehlungen für die Gestaltung der Wissenschaftspolitik auf nationaler und internationaler Ebene auf. Das reichhaltige Informationsmaterial wird durch ein Glossar, eine Auswahl wichtiger Indikatoren sowie ein ausdifferenziertes Material- und ein nach Sachgruppen gegliedertes Literaturverzeichnis erschlossen.

Hinweise für Autoren

Konzept:

Die Zeitschrift „Beiträge zur Hochschulforschung“ bietet Hochschulforschern und Akteuren im Hochschulbereich die Möglichkeit zur Erstveröffentlichung von Artikeln, die wichtige Entwicklungen im Hochschulbereich aus unterschiedlichen methodischen und disziplinären Perspektiven behandeln. Dabei wird ein Gleichgewicht zwischen quantitativen und qualitativen empirischen Analysen, Vergleichsstudien und Überblicksartikeln angestrebt.

Eingereichte Artikel sollten klar und verständlich formuliert, übersichtlich gegliedert sowie an ein Lesepublikum aus unterschiedlichen Disziplinen mit wissenschaftlichem und praxisbezogenem Erwartungshorizont gerichtet sein.

Review-Verfahren:

Wie für eine wissenschaftliche Zeitschrift üblich, durchlaufen alle eingereichten Manuskripte eine zweifache Begutachtung durch anonyme Sachverständige (double blind) innerhalb und außerhalb des Instituts. Dabei kommen je nach Ausrichtung des Artikels folgende Kriterien zum Tragen: Relevanz des Themas, Berücksichtigung des hochschulpolitischen Kontexts, Praxisbezug, theoretische und methodische Fundierung, Qualität der Daten und empirischen Analysen, Berücksichtigung der relevanten Literatur, klare Argumentation und Verständlichkeit für ein interdisziplinäres Publikum. Die Autoren werden über das Ergebnis schriftlich informiert und erhalten gegebenenfalls Hinweise zur Überarbeitung. Die redaktionelle Betreuung der Zeitschrift liegt bei Mitarbeitern des Instituts.

Umfang und Form der eingereichten Manuskripte:

Manuskripte sollten bevorzugt per E-Mail eingereicht werden und einen Umfang von 20 Seiten/50.000 Zeichen mit Leerzeichen nicht überschreiten (Zeilenabstand 1,5, Arial 11). Ergänzend sollten je ein Abstract (maximal 1000 Zeichen mit Leerzeichen) in deutscher und in englischer Sprache sowie Anschrift und Angaben zur beruflichen Funktion des Autors beigefügt sein. Die Druckfassung wird extern von einem Grafiker erstellt. Weitere wichtige Hinweise zur Gestaltung der Manuskripte finden Sie auf unserer Homepage www.ihf.bayern.de unter Publikationen.

Kontakt:

Dr. Lydia Hartwig

Bayerisches Staatsinstitut für Hochschulforschung und Hochschulplanung (IHF)

Prinzregentenstraße 24

80538 München

E-Mail: Hartwig@ihf.bayern.de

Aus dem Inhalt

Sandra Mittag, Rüdiger Mutz, Hans-Dieter Daniel: Anforderungen an Qualitätssicherungsinstrumente für Lehre und Studium an Hochschulen – Ergebnisse einer Meta-Evaluation an der ETH Zürich

Anett Löscher: Qualitätssicherung in Großbritannien durch die Quality Assurance Agency – Ein Überblick

Silke Gülker: Wissenschaft aus Kalkül? Wissenschaftsinterne Antizipation von Wettbewerbsmechanismen im Rahmen von Evaluationsverfahren

Ulf Banscheraus, Mirjam Reiß: Das sächsische Hochschulrating als Beispiel für einen evidenzbasierten Ansatz in der Hochschulentwicklungsplanung – Ein Werkstattbericht