

Beiträge zur HOCHSCHULFORSCHUNG

2 | 2018

Thema: Leistungsbewertung in der Wissenschaft

Mallwitz/Adler/Benzler: Der Einfluss von Evaluierungen auf Forschungseinrichtungen

Kleimann/Hückstädt: Auswahlkriterien in Berufungsverfahren

Behle/Maher: Measuring "Teaching Excellence" in the United Kingdom

Tafertshofer/Werner/Schmidt-Hertha: Reputation von Hochschulstandorten und Studiengängen

Beiträge zur HOCHSCHULFORSCHUNG

2 | 2018

Thema: Leistungsbewertung in der Wissenschaft

Mallwitz/Adler/Benzler: Der Einfluss von Evaluierungen auf Forschungseinrichtungen

Kleimann/Hückstädt: Auswahlkriterien in Berufungsverfahren

Behle/Maher: Measuring "Teaching Excellence" in the United Kingdom

Tafertshofer/Werner/Schmidt-Hertha: Reputation von Hochschulstandorten und Studiengängen

Impressum

Beiträge zur Hochschulforschung

erscheinen viermal im Jahr

ISSN (Print) 0171-645X

ISSN (Online) 2567-8841

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:

Prof. Dr. Katrin Auspurg (Ludwig-Maximilians-Universität München)

Dr. Dr. Lutz Bornmann (Max-Planck-Gesellschaft)

Prof. Dr. Monika Jungbauer-Gans (Universität Hannover, DZHW)

Prof. Dr. Georg Krücken (Universität Kassel, INCHER)

Prof. Dr. Isabell Welpke (Technische Universität München, IHF)

Dr. Lydia Hartwig (IHF)

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:

- Governance von Hochschulen und Forschungseinrichtungen,
- 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,
- Geschlechterungleichheiten in der Wissenschaft,
- Wissenschaft und Wirtschaft,
- International vergleichende Hochschulforschung,
- Wissenschaftsforschung.

Die Zeitschrift veröffentlicht quantitative und qualitative empirische Analysen, Vergleichsstudien, Überblicksartikel und Einblicke in die Praxis, die ein anonymes Peer Review-Verfahren (double blind) durchlaufen haben. Sie bietet die Möglichkeit zum Austausch von Forschungsergebnissen und stellt ein Forum für Hochschulforscher und Experten aus der Praxis dar. Neben Ausgaben, die das gesamte Spektrum der Hochschulforschung abbilden, erscheinen in regelmäßigen Abständen Themenhefte. Hierfür erfolgt in der Regel ein Call for Papers. Manuskripte können jederzeit in deutscher und englischer Sprache eingereicht werden.

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.

Alle Ausgaben der „Beiträge zur Hochschulforschung“ erscheinen in gedruckter Form und werden auf der Homepage unter www.bzh.bayern.de veröffentlicht, die einzelnen Artikel sind nach verschiedenen Kategorien recherchierbar.

Inhalt

Editorial	4
Abstracts	6
Research Note	
Michelle Mallwitz, Philipp Adler, Guido Benzler: Der Einfluss von Evaluierungen auf Forschungseinrichtungen am Beispiel der Leibniz-Gemeinschaft	8
Forschungsartikel	
Bernd Kleimann, Malte Hückstädt: Auswahlkriterien in Berufungsverfahren: Universitäten und Fachhochschulen im Vergleich	20
Heike Behle, Sunil Maher: Measuring “Teaching Excellence” and “Learning Gain” in the United Kingdom	48
Lorenz Tafertshofer, Evamaria Werner, Bernhard Schmidt-Hertha: Grundlagen der Reputation von Studienstandorten: Bewertungsmaßstäbe für die Qualität von Hochschulstandorten und Studiengängen aus der Sicht von sozialwissenschaftlichen Professorinnen und Professoren	68
Buchvorstellungen	90
Hinweise für Autorinnen und Autoren	91

Editorial

Mit den Reform- und Veränderungsprozessen, die das Hochschul- und Wissenschaftssystem seit den 1990er Jahren durchläuft, rückte das Thema „Leistungsbewertung in der Wissenschaft“ zunehmend in den Fokus von Hochschulforschung, Hochschulpolitik und Hochschulmanagement. Insbesondere der Wechsel von einem inputorientierten hin zu einem outputorientierten Steuerungsansatz der Wissenschaft, der auf leistungsorientierte Anreizsysteme zielt, steht im Mittelpunkt vieler Debatten.

Vor diesem Hintergrund widmen sich die „Beiträge zur Hochschulforschung“ in der vorliegenden Ausgabe den unterschiedlichen Facetten des Themas „Leistungsbewertung in der Wissenschaft“:

Michelle Mallwitz, Philipp Adler und Guido Benzler befassen sich mit dem Einfluss von Evaluierungen auf Forschungseinrichtungen am Beispiel der Institute der Leibniz-Gemeinschaft, die sich alle sieben Jahre einer Begutachtung ihrer inhaltlichen und strukturellen Entwicklung stellen müssen. Auf Basis einer Inhaltsanalyse der Förderempfehlungen des Leibniz-Senats kommen die Autoren zu dem Ergebnis, dass neben quantitativ messbaren Leistungen wie DFG-Drittmitteln und Publikationen die Profilbildung eine wichtige Rolle für die Bewertung spielt.

Bernd Kleimann und Malte Hückstädt untersuchen, welche Auswahlkriterien in Berufungsverfahren an Universitäten und welche an Fachhochschulen bevorzugt Verwendung finden. Ihre Analyse ergibt, dass an den Universitäten forschungsbezogene Kriterien dominieren (wissenschaftliche Publikationen, Forschungsprojekte, Drittmittel etc.), während an den Fachhochschulen überdurchschnittlich häufig lehrbezogene Kriterien herangezogen werden (Lehrerfahrung, Ergebnisse von Lehrevaluationen, Lehrkonzept etc.).

Heike Behle und Sunil Maher befassen sich mit dem im britischen Hochschulsystem eingesetzten „Teaching Excellence Framework“ zur Evaluation der Qualität der Lehre, welches die Leistungen der Universitäten hinsichtlich Lehrqualität, Lernumgebung sowie Lernergebnissen und -erträgen quantitativ und qualitativ zu erfassen sucht. Die Autoren diskutieren, ob und wie diese Qualitätsindikatoren adäquat abgebildet werden können, zumal im Hochschulbereich bislang keine international akzeptierten Standards zur Messung von Kompetenzen existieren.

Lorenz Tafertshofer, Evamaria Werner und Bernhard Schmidt-Hertha untersuchen die Reputation von Hochschulstandorten und Studiengängen in den Sozialwissenschaften. Auf Basis einer Onlinebefragung von Professorinnen und Professoren kommen die Autoren zu dem Schluss, dass diese sich vor allem durch die Forschungsreputation sowie die Infrastruktur am Studienort erklärt und andere wichtige Aspekte wie die Qualität der Lehre in den Hintergrund gerückt werden.

Nicholas Folger, Jutta Stumpf-Wollersheim, Isabell Welpke

Abstracts

Michelle Mallwitz, Philipp Adler, Guido Benzler: How institutional evaluations influence research institutions using the example of the Leibniz Association

Evaluations are a constituent part in academia. Institutional evaluations are more and more used as a management tool following the guidelines of New Public Management. Research institutes within the Leibniz Association for example are evaluated every seven years regarding their content and structure. The result of the evaluation decides over further funding of these research institutes. Based on content analyses, the authors examine the criteria that seem to have a major influence in the evaluation process and how these criteria have changed over the course of two evaluation cycles (2003–2007 and 2008–2015). They discuss promising questions that arise from the study, thereby contributing to the discussion of evaluations as management tools.

Bernd Kleimann, Malte Hückstädt: Selection criteria in appointment procedures for professorships: a comparison of universities and universities of applied sciences

The recruiting process of professors in German higher education institutions is based on various selection criteria mirroring the universities' expectations future professors have to meet. Against the backdrop of institutional changes within the German higher education system the question arises which criteria are used by universities and which by universities of applied sciences and to what extent both sets of criteria coincide or differ. The article answers these questions on the basis of empirical data generated by the research project LiBerTas (performance evaluation in appointment procedures – the changing tradition of hiring professors in academia). The authors use multiple correspondence analysis to deconstruct answers of heads of appointment committees concerning the relevance of criteria which are deployed to rank the candidates for a professorship.

Heike Behle, Sunil Maher: Measuring “Teaching Excellence” and “Learning Gain” in the United Kingdom

In the United Kingdom's (UK) higher education system, teaching and learning are currently under review to evaluate both social and individual investment to tertiary education. The UK government introduced the Teaching Excellence Framework (TEF) to assess the quality of undergraduate teaching in universities and other higher education providers. There are 14 research projects which are identifying multiple ways to measure Learning Gain, a key component of the TEF. Teaching Excellence, in the

context of the TEF, encompasses Teaching Quality, the Learning Environment as well as Student Outcomes and Learning Gain, it uses both core metrics and a narrative to evaluate individual universities performance. Current research evaluate methodologies to measure Learning Gain and determine the potential suitability and scalability of possible measures across the sector. This paper describes the current state of policy and research against the specific background of the UK higher education sector.

Lorenz Tafertshofer, Evamaria Werner, Bernhard Schmidt-Hertha: Basics of a study of the reputation of university locations: Evaluation criteria for the quality of university locations and study programs from social science professors' perspective

This paper explores which criteria and which information are central for professors in the social sciences for evaluating university locations and study programs and how far this kind of reputation can be a strong indicator for the quality of study programs. With data from an online survey among German professors in five social science disciplines (n=1 313) it can be shown that the reputation of a location or a study program depends above all on the reputation of colleagues teaching there and the university locations infrastructure. At the same time, significant differences between the subjects involved become visible. Overall, it seems that reputation of a university location or study program gives only poor information regarding the quality of local higher education programs, but that reputation tends to reproduce research indicators.

Der Einfluss von Evaluierungen auf Forschungseinrichtungen am Beispiel der Leibniz-Gemeinschaft

Michelle Mallwitz, Philipp Adler, Guido Benzler

Begutachtungen sind fester Bestandteil des Wissenschaftssystems. In Form von institutionellen Evaluationen werden sie zudem immer stärker als Steuerungsinstrument im Sinne des New Public Managements eingesetzt. Die Forschungseinrichtungen der Leibniz-Gemeinschaft beispielsweise müssen sich spätestens alle sieben Jahre einer Begutachtung ihrer inhaltlichen und strukturellen Entwicklung stellen. Die Beurteilung entscheidet über die Weiterförderung der Einrichtung. Im vorliegenden Beitrag wird auf Basis einer Inhaltsanalyse dargestellt, welche Aspekte in der Bewertung besondere Beachtung finden und wie sich dies im Verlauf der letzten zwei Evaluationsrunden (2003–2007 und 2008–2015) verändert hat. Hiermit soll ein Beitrag zur Diskussion über Evaluation als Steuerungsinstrument geleistet und diskutiert werden, welche weiterführenden Fragen sich am Beispiel der WGL-Einrichtungen ergeben.

1 Einführung

Institutionelle Evaluationen spielen eine immer stärkere Rolle als Steuerungsinstrument an außeruniversitären Forschungseinrichtungen (vgl. u. a. *Heinze 2002*). Die Institute der Leibniz-Gemeinschaft (WGL) etwa müssen sich spätestens alle sieben Jahre einer Begutachtung ihrer inhaltlichen und strukturellen Entwicklung stellen. Die Beurteilung entscheidet über die finanzielle Weiterförderung. Bei erstmaliger negativer Beurteilung folgt die nächste Evaluierung in einem kürzeren Zeitabstand. Fällt diese ebenfalls negativ aus, droht der Einrichtung das Ende der Finanzierung. Der Evaluation kommt somit eine existenzielle Rolle zu.

Im vorliegenden Beitrag wird am Beispiel der WGL-Einrichtungen herausgearbeitet, welche Aspekte in der Evaluierung eine besondere Relevanz für die Überprüfung der wissenschaftlichen Exzellenz und der gesellschaftlichen Relevanz haben und wie sich diese im Verlauf der letzten zwei Evaluationsrunden verändert haben. Als Basis dient eine Inhaltsanalyse der Förderempfehlung des Senats in der veröffentlichten Senatsstellungnahme aller WGL-Einrichtungen, die in zwei aufeinanderfolgenden Evaluationsrunden (2003–2007 und 2008–2015) begutachtet wurden. Auf Basis dieser ersten Annäherung werden weiterführende Fragen hinsichtlich der eingesetzten Kriterien in

der Evaluation außeruniversitärer Forschungseinrichtungen diskutiert und der weitere Forschungsbedarf aufgezeigt. Bewusst ausgespart wird eine normative Einschätzung von Evaluationen als Steuerungsinstrument (siehe hierzu u.a. *Heinze/Arnold 2008, S. 711ff.*).

2 Institutionelle Evaluation

Angesichts dessen, dass der wissenschaftliche Erkenntnisgewinn einerseits hochspezifisch, andererseits von Unsicherheit geprägt ist, sind Peer-Review-Verfahren als Modi der Bewertung dem Wissenschaftssystem inhärent. *Evaluationen* definiert *Wilhelm Krull (2011)* als Betrachtung von Institutionen und Forschungsprogrammen auf einer „höher aggregierten Ebene“ (ebd., S. 15). *Institutionelle Evaluationen* lassen sich als Instrument im Sinne des New Public Managements (NPM) verstehen. NPM bezeichnet die Tendenz, „[ö]ffentliche Dienstleistungen [...] mit Blick auf die Arbeitsorganisation effizienter, hinsichtlich des Preis-Leistungs-Verhältnisses ökonomischer und mit Blick auf die gesellschaftliche Wirkung effektiver zu erbringen“ (*Lange 2008, S. 239*). Es wird ein „Quasi-Markt“ (ebd.) geschaffen, der den Wettbewerb und – so die Annahme – die Leistung der Forschungseinrichtungen stärken soll. Im Zuge des NPM werden staatliche Regulierungen zurückgefahren und die Eigenverwaltung der Forschungsinstitute gestärkt. Gleichzeitig wird die Mittelvergabe an wettbewerbliche Bedingungen geknüpft, und externe Kontrollgremien erhalten größere Bedeutung. Mit einer stärkeren Wettbewerbsorientierung fordern „Politik und Gesellschaft Rechenschaft über Leistungen, Qualität und Nutzen staatlich geförderter Wissenschaft und Forschung“ ein (*Heinze 2002, S. 14*). Rechtfertigung und Legitimierung werden zu zentralen Elementen. Die institutionelle Evaluation als Betrachtung einer gesamten Einrichtung dient hierfür als Messinstrument und betrachtet (vermeintlich) messbare Indikatoren (vgl. u.a. *Kromrey 2003*). Die Evaluation von Forschungsleistungen beruht in diesem Sinne auf dem Verständnis eines „Input-Throughput-Output-Modells“ (*Heinze 2002, S. 15*) von Forschung, in dem Ressourcen eingesetzt werden und am Ende eine Leistung steht.

Kritik am Evaluationsverfahren setzt häufig dort an, wo eine (zu) starke Output-Orientierung in der Betrachtung vermutet wird. Exemplarisch soll hier auf *Margit Osterloh (2010)* verwiesen werden, die auf die Heterogenität unterschiedlicher Disziplinen hinweist, auf Studien zur Subjektivität vermeintlich objektiver Peer-Review-Verfahren sowie auf die Gefahr, nicht intendierte, negative Effekte zu generieren, beispielsweise indem sich die Forscherinnen und Forscher stärker auf die Indikatoren als auf die eigentliche Leistung ausrichten (siehe hierzu u.a. *DFG 2010*).

3 Forschungsgegenstand: Leibniz-Gemeinschaft

Die Leibniz-Gemeinschaft (WGL) wurde 1997 gegründet und geht auf die *Einrichtungen der Blauen Liste*¹ zurück. Aktuell setzt sie sich aus 91 selbstständigen Forschungseinrichtungen zusammen, die von den Natur-, Ingenieur- und Umweltwissenschaften über die Wirtschafts-, Raum- und Sozialwissenschaften bis zu den Geisteswissenschaften reichen. Die WGL hat etwa 18 700 Beschäftigte, darunter 9 500 Wissenschaftlerinnen und Wissenschaftler, und einen gemeinsam von Bund und Ländern finanzierten Etat von insgesamt etwa 1,8 Milliarden Euro (vgl. *Gülker/Simon 2011, S. 86ff.*; *WGL 2017a*). Die Einrichtungen werden in der Regel alle sieben Jahre evaluiert. Ziel dieser Evaluation ist es „zu überprüfen, ob die Voraussetzungen der gemeinsamen Förderung eines Leibniz-Instituts (überregionale Bedeutung und gesamtstaatliches wissenschaftspolitisches Interesse) weiterhin gegeben sind“ (vgl. *WGL 2016*).

Als ehemalige *Einrichtungen der Blauen Liste* haben die WGL-Institute seit den 1980ern eine für deutsche Verhältnisse lange Tradition in der institutionellen Evaluation (vgl. *Krull 2011, S. 21*). 1994 beauftragte die Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung den Wissenschaftsrat (WR) mit einer Systemevaluation der *Einrichtungen der Blauen Liste*. Die Systemevaluation wurde 2000 mit der Stellungnahme des WR abgeschlossen. Die Empfehlungen dieser Stellungnahme bilden die Grundlage für das heutige Evaluationsverfahren (vgl. *WR 2000*).

Die Verantwortung für das Evaluationsverfahren liegt beim Senat der WGL, bestehend aus den „für die gemeinsame Forschungsförderung zuständige[n] Bundes- und Landesminister[n], Präsidenten und Vorsitzende[n] deutscher Wissenschaftsorganisationen sowie weitere[n] Wahlmitglieder[n]“ (*WGL 2017b*). Der Senatsausschuss Evaluierung der WGL stellt für die Evaluation ein Gremium aus Fachgutachterinnen und Fachgutachtern zusammen. Dieses erhält zunächst einen umfangreichen Selbstbericht der Einrichtungen. Es folgt eine Begehung der Einrichtung, an der neben der Gutachtergruppe auch Vertreterinnen und Vertreter von Bund und Land beteiligt sind. Sowohl der Inhalt des Selbstberichts als auch die Kriterien für die Evaluierung sind in den *Grundsätzen des Evaluierungsverfahrens des Senats der Leibniz-Gemeinschaft* geregelt (vgl. *WGL-Senat 2014*).

Die Ergebnisse der Evaluation werden in Form eines Berichts der Einrichtung und dem Zuwendungsgeber zugeleitet (vgl. *Gülker/Simon 2011, S. 84ff.*). Der WGL-Senat erarbeitet auf Basis des Selbstberichts und der Begehung eine wissenschaftspolitische Stellungnahme, die eine Förderempfehlung beinhaltet. Eine erstmalige negative

¹Der Namen geht auf eine „in ihrer ursprünglichen Form auf blauem Papier gedruckten Anlage zur Ausführungsvereinbarung Forschungseinrichtungen der Rahmenvereinbarung Forschung (RV-Fo) aus dem Jahr 1975“ (*Heinz u. Arnold 2008, S. 291*) zurück.

Evaluation führt zu einem verkürzten Zyklus von etwa dreieinhalb Jahren bis zur nächsten Begehung. Eine zweimalige negative Evaluation kann zur Schließung einer Einrichtung führen, „wenn die Qualität der wissenschaftlichen Arbeit nicht gegeben ist, und/oder das gesamtstaatliche Interesse an einem Forschungsgebiet nicht mehr besteht“ (WR 2000, S. 36). Der Druck auf die Einrichtungen, ihre wissenschaftliche Exzellenz und Relevanz nachzuweisen, ist auch deshalb hoch, weil Landesinstitute sich Chancen ausrechnen, in die WGL und damit eine gemeinsame Bund-Länder-Förderung aufgenommen zu werden. Dieser „Aufstieg“ von Landesinstituten ist wiederum nur möglich, wenn er mit einem gleichzeitigen „Abstieg“ bestehender WGL-Einrichtungen einhergeht. Von den im Jahr 2016 veröffentlichten zehn Evaluierungen wurde bei acht Instituten zur vollen Förderdauer geraten. Ein weiteres Institut allerdings wird bereits 2020 wieder evaluiert und eines der Institute wird nicht mehr gefördert (WGL 2016).

4 Methodisches Vorgehen

Grundlage für die Untersuchung sind die über die WGL-Homepage veröffentlichten Förderempfehlungen innerhalb der Senatsstellungen der beiden bisherigen Evaluationszyklen (2003–2007 und 2008–2015).² Insgesamt sind 52 Institute der WGL in die Auswertung eingeflossen, die in beiden bisherigen Evaluationszyklen betrachtet wurden. In einem explorativen Vorgehen wurde in einem ersten Schritt mittels einer qualitativen Inhaltsanalyse herausgearbeitet, welche wiederkehrenden Empfehlungen zur Weiterentwicklung der Einrichtung über die veröffentlichten 104 Förderempfehlungen in der Betrachtung sind (vgl. Mayring 2010). Ausgewertet wurden Textstellen, die Kritik üben oder Verbesserungen nahelegen. In einem zweiten Schritt wurde untersucht, wie häufig die herausgearbeiteten Empfehlungen vorkommen, wobei zwischen der ersten und der zweiten Förderrunde differenziert wurde, um diese Werte schließlich vergleichen zu können. Einschränkend sei darauf hingewiesen, dass das Untersuchungsdesign lediglich die Häufigkeit der Codes betrachtet und damit implizit die Hypothese trifft, dass dies Aussagen über die Relevanz zulässt.

5 Darstellung der Ergebnisse

Im ersten Schritt der Analyse konnten zwölf Codes erhoben werden, die in den Senatsstellungen besonders häufig genannt werden. Als Orientierung dienen dabei die Bewertungskriterien aus den *Grundsätzen des Evaluierungsverfahrens des Senats der Leibniz-Gemeinschaft* (vgl. WGL-Senat 2014). Diese zwölf Codes konnten im Zuge

²Die Senatsstellungen sind unter <https://www.leibniz-gemeinschaft.de/ueber-uns/evaluierung/das-evaluierungsverfahren-des-senats/senatsstellungen/> (Zugriff: 01.10.2017) zugänglich. Im Folgenden wird auf die jeweilige Stellungnahme unter Angabe des Kürzels der Einrichtung und des jeweiligen Jahres Bezug genommen.

einer Clusteranalyse zu vier Kategorien zusammengefasst werden. Die Kategorie *Gesamtkonzept und Profil* greift die strategische Ausrichtung auf, während die anderen drei Kategorien *Publikationsleistungen*, *Drittmittleinnahmen* und *Transfer* die erbrachten Leistungen der vergangenen drei Jahre in Betracht ziehen.

Die nachstehende Tabelle zeigt die entsprechenden Codes und Kategorien mit beispielhaften Ankerzitaten (vgl. Tabellen 1–4).

Tabelle 1: Codes und Ankerzitate Gesamtkonzept und Profil

Gesamtkonzept und Profil	
Code	Ankerzitat
Profilbildung	„Auf der Grundlage seiner Themen, Methoden und Infrastrukturen hat das ZALF ein Gesamtkonzept erarbeitet, dessen Kohärenz und Fokussierung zukünftig weiter gesteigert werden muss.“ (ZALF 2013, S. 2)
Strategische Weiterentwicklung	„Damit die GESIS für die Zukunft gerüstet ist, müssen die methodische Forschung gestärkt und vor allem integrierte Serviceangebote adressatengerecht mit der sozialwissenschaftlichen Klientel entwickelt werden.“ (GESIS 2005, S. 4)
Institutsübergreifende Zusammenarbeit	„Nach wie vor ist eine Verbesserung auf der Ebene der Abteilungs- und Forschungsgruppenleitungen nötig, um das gemeinsame Forschungspotential am HPI so weit wie möglich ausschöpfen zu können.“ (HPI 2010, S. 3)

Tabelle 2: Codes und Ankerzitate Publikationsleistungen

Publikationsleistungen	
Code	Ankerzitat
Quantitative Steigerung	„Kritisch ist vor allem, dass ein Teil der Wissenschaftler wenig oder gar nicht publiziert.“ (ATB 2007, S. 3)
Qualitative Steigerung	„Insbesondere die Anzahl der Veröffentlichungen in referierten Zeitschriften [...] sollten weiter gesteigert werden.“ (IfW 2005, S. 3)
Publikationsstrategie	„Empfehlenswert sind die Erarbeitung eines Publikationskonzeptes [...]“ (FIS 2006, S. 3)

Tabelle 3: Codes und Ankerzitate Drittmittleinnahmen

Drittmittel	
Code	Ankerzitat
Steigerung der gesamten Drittmittel	„Der Senat erwartet, dass das ZPID die Drittmittelerträge in den kommenden Jahren im Rahmen des Ausbaus von Forschung und Entwicklung deutlich steigert.“ (ZPID 2011, S. 4)
Steigerung der DFG-Mittel	„Allerdings empfiehlt der Senat, den Anteil von DFG-Mitteln innerhalb des Drittmittelportfolios, zu dem derzeit nur einzelne Institute und Beschäftigte beitragen, erheblich zu steigern.“ (ZALF 2013, S. 3)
Steigerung der EU-Mittel	„Insbesondere die Einnahmen aus Drittmitteln von der DFG und der EU muss das ATB deutlich erhöhen.“ (ATB 2014, S. 3)

Tabelle 4: Codes und Ankerzitate Transfer

Transfer	
Code	Ankerzitat
Qualität der Transferleistungen	„Hinsichtlich der Konjunkturanalysen und der Branchenforschung hat die Bewertungsgruppe festgestellt, dass sich die Qualität der zugrundeliegenden ökonomischen Modelle, statistischen Ansätze und Umfragemethoden stärker an dem heutigen Stand der Forschung orientieren sollte.“ (ifo 2006, S. 3)
Kontakte zu Politik und Gesellschaft	„Darüber hinaus wäre es wünschenswert, wenn das IDS versuchte, auf die Gestaltung des Urheberrechts mit dem Ziel einzuwirken, der Wissenschaft den Zugriff auf KobRA zu Forschungszwecken zu ermöglichen.“ (IDS 2004, S. 3)
Patente und Ausgründungen	„Der Senat empfiehlt, die einzelnen Aktivitäten zur Verwertung von Arbeitsergebnissen in einer übergreifenden Verwertungsstrategie zusammenzuführen.“ (IKZ 2012, S. 3)

Im zweiten Schritt wurde eine Zählung dieser Bewertungsaspekte vorgenommen und beide Evaluationsrunden einander gegenübergestellt. Eine Betrachtung der Nennungen der Schwachstellen und Kritikpunkte ohne eine vorherige Aufteilung auf die beiden Evaluationsperioden verdeutlicht, dass der Bereich *Drittmittelnahmen* am häufigsten genannt wird (123 Nennungen). An zweiter und dritter Stelle stehen *Publikationsleistungen* (106 Nennungen) beziehungsweise die Aspekte rund um das *Gesamtkonzept und Profil* (97 Nennungen). *Transfer* spielt mit 43 Nennungen für die Evaluationen laut den vorliegenden Ergebnissen eine eher untergeordnete Rolle (vgl. Tabelle 5).

Tabelle 5: Anzahl der Nennungen nach Codes

Kategorie	Codes	Anzahl der Nennungen	Anzahl der Nennungen
Gesamtkonzept und Profil	Profilbildung	56	97
	Strategische Weiterentwicklung	18	
	Institutsübergreifende Zusammenarbeit	23	
Publikationsleistung	Quantitative Steigerung	46	106
	Qualitative Steigerung	54	
	Publikationsstrategie	6	
Drittmittelnahmen	Steigerung	39	123
	DFG-Mittel	51	
	EU-Mittel	33	
Transfer	Qualität der Transferleistungen	23	43
	Kontakte zu Politik und Gesellschaft	11	
	Patente und Ausgründungen	9	
		369	

Bei genauerer Betrachtung der einzelnen Themenbereiche erscheinen manche Kritikpunkte besonders entscheidend. So wurden beispielsweise die Profilbildung, die Steigerung der DFG-Mittel und die qualitative Steigerung der Publikationen jeweils etwa bei der Hälfte der untersuchten Evaluationen als Kritikpunkt genannt. Ebenfalls von Bedeutung für viele der Einrichtungen aus Sicht der Förderempfehlung ist die quantitative Steigerung von Publikationsleistungen. Betrachtet nach Evaluationsrunden ist die Profilbildung das meistgenannte Kriterium in den Empfehlungen.

Tabelle 6: Anzahl der Nennungen nach Evaluationsrunden (2003–2007, 2008–2015)

Kategorie	Codes	Anzahl der Nennungen (2003–2007)	Anzahl der Nennungen (2008–2015)
Gesamtkonzept und Profil	Profilbildung	32	24
	Strategische Weiterentwicklung	10	8
	Institutsübergreifende Zusammenarbeit	17	6
Publikationsleistung	Quantitative Steigerung	27	19
	Qualitative Steigerung	30	24
	Publikationsstrategie	2	4
Drittmittelannahmen	Steigerung	22	17
	DFG-Mittel	29	22
	EU-Mittel	20	13
Transfer	Qualität der Transferleistungen	13	10
	Kontakte zu Politik und Gesellschaft	4	7
	Patente und Ausgründungen	6	3
		212	157

Insgesamt ist die Zahl der kritischen Nennungen der verschiedenen Aspekte zwischen den beiden Bewertungsperioden zurückgegangen (vgl. Tabelle 6). Besonders auffällig ist dies bei den Punkten *institutsübergreifende Zusammenarbeit* (17 Nennungen → 6 Nennungen), *quantitative Steigerung der Publikationsleistungen* (27 Nennungen → 19 Nennungen) und *Steigerung von DFG- und EU-Mitteln* (jeweils sieben Nennungen weniger als in der vorherigen Evaluationsrunde). Nur zwei Codes wurden in der zweiten Evaluationsrunde häufiger genannt als in der ersten. Dabei ist auffällig, dass es sich im Verhältnis zu den Gesamtnennungen um eher weniger relevante Punkte zu handeln scheint (*Kontakte zu Politik und Gesellschaft* und *Publikationsstrategie*). Überraschend scheint, dass die Aspekte in der Kategorie *Transfer* einen so geringen Stellenwert einnehmen, obwohl der Austausch mit Gesellschaft, Politik und Wirtschaft zum Leitbild der WGL gehört.

6 Diskussion der Ergebnisse

Die besonders häufige Nennung des Aspekts *Profilbildung* deutet daraufhin, dass neben den quantitativ messbaren Leistungen die strategische Ausrichtung der Einrichtungen eine wichtige Rolle für ihre Bewertung spielt. *Silke Gülker* und *Dagmar Simon* kamen in einer qualitativen Betrachtung der Evaluationen ausgewählter WGL-Einrichtungen zu dem Schluss, dass die institutionellen Evaluationen eine Reflektion der strategischen Ausrichtung mit sich bringen, da auf inhaltlicher Ebene eine Einschätzung der ganzen Einrichtung vorgenommen wird (ders. 2011, S. 90). Die häufige Bezugnahme auf das Thema *Profilbildung* in der Analyse scheint dies zu stützen. *Heinze* und *Arnold* (2008) weisen in diesem Zusammenhang darauf hin, dass die WGL gegenüber den anderen großen außeruniversitären Forschungseinrichtungen aufgrund der Zusammenstellung und der internen Konkurrenzsituation eine weitaus größere Binnendifferenzierung aufweist. (vgl. ebd.) Die Profilierung erfolgt damit stärker auf Einrichtungsebene als auf Ebene der Gesamtorganisation, was diesem Merkmal zusätzlich Gewicht verleiht.

Die Bedeutung quantitativ messbarer Daten, die *Gülker* und *Simon* (2011) noch als sehr hoch einschätzten, scheint in der vorliegenden Betrachtung in der zweiten Evaluationsrunde leicht abzunehmen. Es zeigt sich, dass innerhalb der Evaluation als NPM-Instrument weitere Instrumente im Sinne des NPM gefördert werden. Deutlich macht dies beispielsweise dieser Auszug aus einer Senatsstellungnahme: *„Der Senat empfiehlt den Zuwendungsgebern, dem Institut nach Einführung von Kosten-Leistungsrechnung und Programmbudgets nun auch die benötigte Flexibilität bei der Bewirtschaftung der Ressourcen zu gewähren, indem die Stellenpläne flexibilisiert werden und eine überjährige Bewirtschaftung des Haushaltes ermöglicht wird.“* (WGL-Senat 2006, S. 2)

Die Verbindung aus der Betrachtung der Strategie der Einrichtung sowie der Forschungsleistungen kann als Versuch verstanden werden, sowohl die Vergangenheit als auch die Zukunft in die Bewertung miteinzubeziehen. Dabei können sinnvolle und ausformulierte strategische Ziele und ein klares Profil eine hohe Leistungsfähigkeit in den kommenden Jahren zwar nicht garantieren, jedoch ermöglichen sie eine Einschätzung der Ausrichtung und der Ziele der Einrichtung.

Interessant ist, dass der Transfer wissenschaftlicher Erkenntnis – obwohl elementarer Teil des WGL-Selbstverständnisses – in der Evaluation eine untergeordnete Rolle zu spielen scheint. Hier spielt es sicherlich eine Rolle, dass die außeruniversitären Forschungseinrichtungen – trotz größerer Ressourcenausstattung pro Professorin/Professor – gegenüber den Universitäten *„nur etwa ein Fünftel zur Gesamtmenge aller vom staatlich finanzierten Forschungssektor produzierten wissenschaftlichen*

Publikationen“ beitragen (Heinze und Arnold, 2008, S. 693) und damit der Fokus in den Empfehlungen stärker auf den Publikationsleistungen als auf dem Transfer liegt.

Die Ergebnisse zeigen, dass die Nennung kritischer Aspekte in den Senatsstellungen im zweiten Durchlauf insgesamt abnimmt. Für den Rückgang können zwei mögliche Erklärungsansätze herangezogen werden: Es ist möglich, dass die Institute auf die Empfehlungen der ersten Runde der Evaluierungen reagiert und Veränderungen angestrebt haben. Diese Hypothese könnte in einer weiterführenden Untersuchung mit Blick auf die Selbstberichte der Einrichtung in der zweiten Evaluationsrunde untersucht werden. Heinze und Arnold (2008) weisen nach, dass die WGL sich einem „seitens der Forschungspolitik stark gewachsenen Druck nach effizientem Ressourceneinsatz und Anwendungsorientierung“ (ebd., S. 711) stellen muss, im Zuge dessen die WGL-Einrichtungen deutlich höhere Zuwachsraten bei Publikationen und Patentmeldungen verzeichnen als beispielsweise die Max-Planck- oder die Fraunhofer-Gesellschaft (vgl. ebd.). Gülker und Simon (2011) weisen darauf hin, dass Empfehlungen aus der Evaluation heraus entweder direkt umgesetzt oder aber die entsprechenden Daten im Selbstbericht in der nächsten Runde positiver präsentiert werden (vgl. ebd., S. 91). Stefan Hornbostel (2011) spricht in diesem Zusammenhang zynisch von „mechanistischen Umsetzungen“ und „ressourcenintensive(n) Inszenierungen“ (ebd., S. 9). Auffällig ist dennoch, dass in beiden Runden genau gleich viele Institute mit einer verkürzten Förderdauer belegt und damit „schlecht“ bewertet wurden.

Eine zweite Schlussfolgerung könnte sein, dass sich aufgrund übergeordneter wissenschaftspolitischer Diskurse der Fokus des Senats gewandelt hat und nun neben den Leistungen auch die weniger gut quantitativ abbildbaren Aspekte wie Governance der Einrichtung oder die Vernetzung mit anderen Instituten oder Hochschulen stärker in den Blick genommen werden. Hier stellt sich also die Frage, inwieweit die Einrichtungen sich in der zweiten Runde – gemessen an den in den Förderempfehlungen gewählten Kriterien – tatsächlich verbessert haben oder inwieweit sie gelernt haben, sich besser zu präsentieren.

Es ist zu vermuten, dass eine reduzierte Förderdauer höheren Veränderungsdruck erzeugt. Gülker und Simon (2011) deuten in ihrer Untersuchung darauf hin, dass die institutionelle Evaluation der Einrichtungsleitung ebendieser eine neue Möglichkeit der Steuerung und der Legitimation ihres Handelns gibt (vgl. ebd., S. 91). In beiden Evaluationsrunden ist der Fokus auf das Profil und die Weiterentwicklung der Einrichtung hoch. Aspekte dieses Clusters beziehen sich ganz wesentlich auf die Aufgabenbereiche der Einrichtungsleitung. Hinzu kommt, dass im Rahmen der Senatsstellungen immer wieder auch Punkte genannt werden, die als NPM-Instrumente die Leitung der Einrichtung stärken.

7 **Ausblick**

Die Diskussion der Ergebnisse bietet im Sinne der Research Note eine Basis für weiterführende Hypothesen: Dem Vorwurf, dass Evaluationen keinen Einfluss hätten, kann auf Basis der vorliegenden Erhebung widersprochen werden. Ob allerdings die gewünschten Wirkungen erzielt werden oder die Einrichtungen beispielsweise „nur“ ihre Selbstinszenierungsfähigkeiten trainieren oder aber andere, nicht intendierte Wirkungen mit der Evaluation einhergehen, kann diese Untersuchung nicht beantworten. *Torka (2011)* weist generell auf die Schwierigkeit hin, kausale Mechanismen zwischen Evaluationen und wissenschaftlichem Handeln empirisch zu untersuchen (vgl. ebd. S. 70).

Aufbauend auf dieser ersten Betrachtung schließen sich daher weiterführende interessante Fragen an: In Hinblick auf die quantifizierbaren Kriterien wie etwa die Publikationsleistungen ließe sich beispielsweise untersuchen, ob sich zwischen Einrichtungen, die positiv bewertet wurden und solchen, die mit einer kürzeren Frist bis zur nächsten Evaluation belegt wurden, Unterschiede zeigen. Durch die stete Zunahme der Datengrundlage in Form von zusätzlichen Evaluierungen samt Evaluierungstexten sowie die immer transparentere Aufbereitung von Publikationsleistungen ist dies ein Ansatz, der sich auf Basis der guten Datengrundlage mittelfristig weiterverfolgen lässt. Erste statistische Auswertungen zeigen durchaus eine Veränderung der Publikationsleistungen nach negativen Evaluationen auf, doch ermöglicht die geringe Zahl der negativen Evaluierungen noch keine signifikante Aussage.

Auch wäre es für die weitere Betrachtung der Evaluation als Steuerungsinstrument fruchtbar zu untersuchen, inwieweit die Empfehlungen der Senatsstellungen sich auf die Position der Einrichtungsleitung auswirken und wie sie von dieser genutzt werden. Dieser Ansatz könnte mit qualitativen Interviews mit den Leitungen und der mittleren Führungsebene der Einrichtung vorangetrieben werden, auch im Hinblick auf Unterschiede bei sehr positiven und eher negativen Bewertungen.

In der Senatsstellungnahme wurde ein besonderer Fokus auf die Empfehlung des Senats gelegt (in Abgrenzung zum Bericht der unabhängigen Bewertungsgruppe, der neben dem Bericht der evaluierten WGL-Einrichtung ebenfalls Bestandteil der Senatsstellungnahme ist). Diese abschließende Empfehlung greift die Einschätzungen der Fachgutachterinnen und Fachgutachter auf, ist darüber hinaus aber auch aufgrund der Aufgabe und Zusammensetzung des Senats in seiner Einschätzung stärker wissenschaftspolitisch geleitet. In einem weiteren Schritt wäre es daher lohnenswert, die Kriterien und deren Häufigkeit in den Berichten der Fachgutachterinnen und Fachgutachter den abschließenden Empfehlungen des Senats gegenüberzustellen, um zu vergleichen, wie viel Einfluss die Gutachtergruppe tatsächlich auf die finale Entscheidung hatte. Eine Auffälligkeit, die sich bei der Untersuchung des Materials zeigte, ist, dass seitdem sich die DFG für ein Kaskadenmodell in der Frauenförderung aus-

gespröchen hat, dieser Begriff öfter in den Senatsstellungen auftauchte. Interessant könnte es dementsprechend sein, die Codes und Kategorien in den Senatsstellungen mit anderen einflussreichen wissenschaftspolitischen Papieren zu vergleichen und in diesem Kontext zu deuten.

Literatur

DFG (2010): DFG-Pressemitteilung Nr. 7, 23. Februar 2010 http://www.dfg.de/service/presse/pressemitteilungen/2010/pressemitteilung_nr_07/index.html (Zugriff: 04.11.2016)

Gülker, Silke; Simon, Dagmar (2011): Nach der Evaluation ist vor der Evaluation? Institutionelle Folgen von Forschungsbewertungen im internationalen Vergleich. In: Hornbostel, Stefan; Schelling, Anna (Hrsg.): Evaluation: New Balance of Power? iFQ-Working Paper No. 9. Berlin, S. 83–93

Heinze, Thomas; Arnold, Natalie (2008): Governanceregimes im Wandel. Eine Analyse des außeruniversitären, staatlich finanzierten Forschungssektors in Deutschland. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 60, 2008, S. 686–720

Heinze, Thomas (2002): Evaluation von Forschungsleistungen: Konzeptionelle Überlegungen und Situationsbeschreibung für Deutschland. In: Wissenschaftsmanagement, 2, 2002, 6, S. 14–22

Hornbostel, Stefan (2011): Resonanzkatastrophen, Eigenschwingungen, harmonische und chaotische Bewegungen. In: Hornbostel, Stefan; Schelling, Anna (Hrsg.): Evaluation: New Balance of Power? iFQ-Working Paper No. 9. Berlin, S. 7–14

Kromrey, Helmut (2003): Evaluation in Wissenschaft und Gesellschaft. Vortrag bei der Feierstunde zur Eröffnung des Centrums für Evaluation am 21.2.2003 in Saarbrücken. http://www.hkromrey.de/Kromrey_EvalWissG.pdf (Zugriff: 04.11.2016)

Krull, Wilhelm (2011): Bewertung, Begutachtung und Evaluation in Wissenschaft und Forschung. In: Hornbostel, Stefan; Schelling, Anna (Hrsg.): Evaluation: New Balance of Power? iFQ-Working Paper No. 9. Berlin, S. 15–23

Lange, Stefan (2008): *New Public Management und die Governance der Universitäten*. In: dms – der moderne staat – Zeitschrift für Public Policy, Recht und Management, 1, 2008, S. 235–248

Mayring, Philipp (2010): Qualitative Inhaltsanalyse. Grundlagen und Techniken. Weinheim

Osterloh, Margit (2010): *Governance by Numbers*. Does it really work in research? In: Analyse & Kritik, 2, 2010, S. 267–283

Torka, Marc (2011): Institutioneller gleich handlungspraktischer Wandel? Das Beispiel von Begutachtungspraktiken bei der Evaluation wissenschaftlicher Einrichtungen. In: Hornbostel, Stefan; Schelling, Anna (Hrsg.): Evaluation: New Balance of Power? iFQ-Working Paper No. 9. Berlin, S. 69–82

WGL (2017a): Die Leibniz-Gemeinschaft. <http://www.leibniz-gemeinschaft.de/ueberuns/> (Zugriff: 01.10.2017)

WGL (2017b): Senat. <https://www.leibniz-gemeinschaft.de/ueber-uns/organisation/senat/> (Zugriff: 01.10.2017)

WGL (2016): Das Evaluierungsverfahren des Senats der Leibniz-Gemeinschaft. <http://www.leibniz-gemeinschaft.de/ueber-uns/evaluierung/das-evaluierungsverfahren-des-senats/> (Zugriff: 04.11.2016)

WGL-Senat (2014): Grundsätze des Evaluierungsverfahrens des Senats der Leibniz-Gemeinschaft. http://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Evaluierung/Austausch_August_2014/Anlage_3_-_Kriterien_fuer_die_Evaluierung_von_Einrichtungen_der_WGL.pdf (Zugriff: 04.11.2016)

WGL-Senat (2006): Stellungnahme zum Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK). https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/downloads/Evaluierung/Senatsstellungen/Senatsstellungnahme-IPK-2006.pdf (Zugriff: 01.10.2017)

WR (2000): Systemevaluation der Blauen Liste – Stellungnahme des Wissenschaftsrates zum Abschluß der Bewertung der Einrichtungen der Blauen Liste. <http://www.wissenschaftsrat.de/download/archiv/4703-00.pdf> (Zugriff: 04.11.2016)

Artikel eingereicht: 01.10.2017
Artikel angenommen: 09.01.2018

Anschriften der Autoren:

Michelle Mallwitz M.A.
Referentin für Studiengangsentwicklung
Goethe-Universität Frankfurt am Main
Theodor-W.-Adorno-Platz 1
60323 Frankfurt am Main
E-Mail: mallwitz@em.uni-frankfurt.de

Jan Philipp Adler M.Sc.
Berater für Hochschulen und außeruniversitäre Forschungseinrichtungen
rheform EntwicklungsManagement GmbH
Kasernenstraße 27
40213 Düsseldorf
E-Mail: philipp.adler@rheform.de

Prof. Dr. Guido Benzler
Geschäftsführer der rheform EntwicklungsManagement GmbH, Beratungsunternehmen für Hochschulen und außeruniversitären Forschungseinrichtungen
rheform Entwicklungsmanagement GmbH
Kasernenstraße 27
40213 Düsseldorf
E-Mail: guido.benzler@rheform.de

Auswahlkriterien in Berufungsverfahren: Universitäten und Fachhochschulen im Vergleich

Bernd Kleimann, Malte Hückstädt

Bei der Auswahl von Bewerberinnen und Bewerbern für eine Professur im Rahmen von Berufungsverfahren an staatlichen deutschen Hochschulen werden verschiedene Auswahlkriterien eingesetzt, in denen sich die Erwartungen der Hochschulen an die zukünftigen Stelleninhaber niederschlagen. Vor dem Hintergrund der institutionellen Veränderungen im deutschen Hochschulsystem stellt sich dabei die Frage, welche Kriterien an Universitäten und welche an Fachhochschulen¹ Verwendung finden und inwiefern sich beide Kriteriensets gleichen bzw. unterscheiden.

1 Einleitung: Auswahlkriterien in Berufungsverfahren unter den Bedingungen institutionellen Wandels

Für Hochschulen ist die Entscheidung über die Auswahl von Bewerberinnen und Bewerbern für vakante Professuren von zentraler Bedeutung. Bei dieser Entscheidung kommen Auswahlkriterien zum Einsatz, mit denen die Anforderungen der Hochschulen an die potenziellen Stelleninhaberinnen und Stelleninhaber operationalisiert werden. Rückt man nun die institutionelle Differenz von Universitäten und Fachhochschulen in den Mittelpunkt, so stellt sich – insbesondere vor dem Hintergrund der Veränderungsprozesse und Reformen im deutschen Hochschulsystem – die Frage, nach welchen Gesichtspunkten Universitäten und Fachhochschulen heute ihr professorales Personal auswählen und in welchem Verhältnis die Kriterien beider Hochschultypen zueinander stehen.

Diesen Fragen geht der folgende Artikel auf der Basis von Daten des Forschungsprojekts „Leistungsbewertung in Berufungsverfahren“² mithilfe einer multiplen

¹Im Folgenden ist der Kürze halber nur von „Fachhochschulen“ die Rede. Dabei sind Hochschulen für angewandte Wissenschaften stets inbegriffen.

²Das Projekt wurde vom BMBF in der Förderlinie „Leistungsbewertung in der Wissenschaft“ gefördert (FKZ 01PY13008).

Korrespondenzanalyse nach.³ Dabei stützt sich die Annahme, dass die aufgeworfenen Fragestellungen sowohl für die Hochschulforschung als auch für die Hochschulpraxis relevant sind, auf Beobachtungen grundlegender Veränderungsprozesse im deutschen Hochschulsystem und auf die gestiegene Bedeutung von Berufungsverfahren für das strategische Handeln der Hochschulen.

Bezüglich der Entwicklung des deutschen Hochschulsystems lautet eine weithin geteilte Diagnose, dass sich seine Rahmenbedingungen im Zuge der Implementierung des neuen Steuerungsmodells verändert haben (*De Boer/Enders/Schimank 2008; Schimank 2014*).⁴ Das Hochschulsystem unterliegt folglich einem *institutionellen Wandel* (*Walgenbach/Meyer 2008*), das heißt einem Umbruch in Bezug auf gesellschaftlich verankerte Erwartungen, die Zwecksetzung, Organisationsstruktur, Art der Steuerung oder Personalstruktur von Hochschulen betreffen.

Zwei Stränge dieses Wandels sind für die Frage nach den Auswahlkriterien in Berufungsverfahren an Universitäten und Fachhochschulen von Bedeutung: erstens die „Organisations-“ (*Huber 2012; von Wissel 2007*) oder „Akteurwerdung“ (*Krücken/Meier 2006; Meier 2009*) der deutschen Hochschulen und zweitens ein *blurring of boundaries* zwischen den bislang systemprägenden Hochschultypen Universität und Fachhochschule (*Wissenschaftsrat 2010; Duong/Hachmeister/Roessler 2014; Schreiterer 2016*). Während die Akteurwerdung der Hochschulen zur Folge hat, dass ihnen mit steigender Autonomie auch ein größeres Maß an Verantwortung für ihre Angelegenheiten (inklusive der Personalpolitik) zukommt, lässt sich aus der (behaupteten) vertikalen (Ent-)Differenzierung des zweischichtigen deutschen Hochschulsystems die Schlussfolgerung ziehen, dass sich die Auswahlkriterien von Universitäten und Fachhochschulen in Berufungsverfahren angenähert haben müssten. Beide Aspekte des Wandels werden im Folgenden näher beleuchtet.

³Ziel der multiplen Korrespondenzanalyse ist es, eine sowohl theoretisch plausible als auch empirisch fundierte „Raumkarte“ zu erzeugen, die die von den befragten Berufungskommissionsvorsitzenden gewichteten Auswahlkriterien detailliert darstellt. Zum einen werden die forschungsleitenden Annahmen mittels der multiplen Korrespondenzanalyse einem (ersten) „Säurebad der empirischen Prüfung“ (*Roose/Rössel 2015, S. 10*) ausgesetzt. Zum anderen ist es im Allgemeinen nicht angeraten, der multiplen Korrespondenzanalyse (auch wenn sie, wie etwa das klassische Regressionsverfahren, keinem Sparsamkeitsgebot unterliegt) theorie- und also wahllos Variablen-Sets zuzuspielen: „Performing a geometric analysis does not mean gathering disparate data and looking for, what comes out of the computer“ (*Le Roux/Rouanet 2010, S. 10*). Die multiple Korrespondenzanalyse und Theorie werden daher im Folgenden miteinander verschränkt: Aus den theoretischen Annahmen ergibt sich die Auswahl konkreter Variablen(-sets), die entweder aktiv oder passiv der multiplen Korrespondenzanalyse zugespielt werden. Der durch die multiple Korrespondenzanalyse aufgespannte Raum dieser Merkmale wird in einem weiteren Schritt auf die theoretischen Annahmen rückbezogen, um letztere zu verwerfen, zu ergänzen oder zu konkretisieren.

⁴Hierzu zählen: ein verstärkter Wettbewerb und ein gestärktes Management, kontraktualistische Steuerungsformen, der Einfluss neuer intermediärer Einrichtungen wie Akkreditierungsagenturen etc., zugleich aber auch der partielle Rückzug des Staates aus der Detailsteuerung und eine Schwächung der akademischen Selbstverwaltung.

2 Akteurwerdung in Hochschulen

Im Zuge der Akteurwerdung der Hochschulen ist die Verantwortung für das Berufungsgeschehen in den meisten Ländern vom Ministerium abgezogen und auf die Hochschulen selbst verlagert worden (*Klawitter 2017, S. 23*). Die Hochschulen müssen Berufungsverfahren daher eigenverantwortlich gemäß den gesetzlichen Verfahrensanforderungen steuern. Diese Steuerung liegt zugleich im elementaren Eigeninteresse der Hochschulen, da mit der Berufung einer Professorin oder eines Professors in sachlicher, sozialer und zeitlicher Hinsicht weitreichende Festlegungen verbunden sind (*Kleimann/Klawitter 2017*). In *sachlicher* Hinsicht wird das thematische Profil von Hochschule und Fakultät in Forschung und Lehre durch die berufene Person entscheidend mitbestimmt. In *sozialer* Hinsicht wirkt sich die Berufungsentscheidung intern auf die Art der Interaktion zwischen Hochschulmitgliedern und das Klima von Entscheidungsprozessen aus. Externe Effekte ergeben sich durch das in den Netzwerken der Professoren gespeicherte Sozialkapital. In *zeitlicher* Hinsicht bindet sich eine Hochschule durch die Verbeamtung auf Lebenszeit (die heute immer noch der Regelfall ist) auf lange Zeit an die ausgewählte Person. Gesteigert wird die Relevanz von Berufungsentscheidungen zudem dadurch, dass mit der Einführung der W-Besoldung und der Flexibilisierung der Besoldung (vgl. *Preißler 2016*) der Wettbewerb unter den Hochschulen um die „besten Köpfe“ angeheizt worden ist. Schließlich unterliegt die Berufungspraxis dem normativen Druck veränderter gesellschaftlicher Wertvorstellungen – insbesondere mit Blick auf die Gleichstellung der Geschlechter durch die geforderte Erhöhung des Frauenanteils unter den Professoren.⁵

Unter diesen Bedingungen arbeiten viele Hochschulen darauf hin, ihre Berufungspraxis so zu professionalisieren, dass sie den organisationalen und gesellschaftlichen Anforderungen gerecht wird. Zu den Bemühungen um ein Berufsmanagement gehören unter anderem die Verabschiedung von Berufsordnungen und Leitfäden für Berufungskommissionen, die Einrichtung spezieller Stellen und Funktionen (z. B. Berufsbeauftragte, fachfremde Beobachter u. ä.), die Neugestaltung von Ämtern (z. B. Prorektorate oder Vizepräsidenten für Berufung), die Etablierung innovativer Qualitätssicherungsverfahren (z. B. Auswahlgespräche von Listenplatzierten mit Personalberatern, Schulungen für Berufungskommissionsmitglieder) oder die Bereitstellung von Informationsgelegenheiten für die Bewerber (z. B. webbasierte Berufsportale).

Für die hier verfolgte Fragestellung sind diese Entwicklungen insofern von Bedeutung, als die Relevanzsteigerung des Berufungsprozesses erwarten lässt, dass sich das Aufgabenprofil und die Identität der Hochschule sowie ihre Anforderungen an das

⁵In der Online-Umfrage des Projekts „Leistungsbewertung in Berufungsverfahren“ haben über 40 Prozent der Universitätsleitungen und sogar fast 60 Prozent der Fachhochschulleitungen angegeben, dass Gleichstellungsaspekte schon bei der Konstruktion einer Professur (Festlegung von Denomination und Ausstattung) wichtig bis sehr wichtig sind.

(zukünftige) professorale Personal direkt in den eingesetzten Auswahlkriterien niederschlagen.

Die zweite institutionelle Veränderungsdynamik, die die Berufung von Professorinnen und Professoren tangiert, ist die (behauptete) Erosion der Unterscheidung zwischen den Organisationstypen Universität und Fachhochschule. Diese Veränderung wird im Folgenden als „Sog nach oben“ (*Schreiterer 2016, S. 28*), d. h. als Angleichung der Fachhochschulen an die Universitäten gedacht. Ob bzw. in welchem Maße eine solche Assimilation tatsächlich besteht, ist umstritten (vgl. *Wissenschaftsrat 2010; Duong/Hachmeister/Roessler 2014; Schreiterer 2016*). So gehen die Meinungen darüber auseinander, inwieweit es sich bei dieser „Entwicklung“ um eine normative Diskursposition von Akteuren handelt, die an einer Stärkung der Fachhochschulen interessiert sind, oder ob man es mit einem Realphänomen zu tun hat, als dessen erste Anzeichen die Übertragung des Promotionsrechts auf Fachhochschulen in Hessen oder der steigende Anteil der Forschungsausgaben im Fachhochschulsektor angeführt werden.

Die (umstrittene) Annäherung zwischen Universität und Fachhochschule kann dabei als *Ent-* wie auch als *Ausdifferenzierung* beschrieben werden (vgl. *Duong/Hachmeister/Roessler 2014, S. 9ff.*) Eine Entdifferenzierung lässt sich darin erkennen, dass die strukturellen Unterschiede zwischen Universität und Fachhochschule reformbedingt immer stärker verblassen und sich dadurch die Zweischichtigkeit des Systems schrittweise auflöst. Andererseits kann man diesen Prozess zugleich als institutionelle *Ausdifferenzierung* des Gesamtsystems charakterisieren (*Schnell 2013*). Demnach ist nicht mehr die Koexistenz zweier Hochschultypen, sondern die allmähliche Ausbildung eines Spektrums neuer Hochschulformen systemprägend (*Wissenschaftsrat 2010*). Welche Beschreibung man vorzieht, ist für die hier relevante Kontroverse allerdings zweitrangig: Letztere kreist ja um die grundsätzlichere Frage, ob es eine Annäherung zwischen Universität und Fachhochschule – sei es als *Ent-*, sei es als *Ausdifferenzierung* – überhaupt gibt.

3 Differenz- versus Konvergenzthese

Die skizzierte Kontroverse kann im Kontext der hier verfolgten Fragestellung genutzt werden, um (zugespitzte) theoretisch begründete Erwartungen an die Konfigurationen konkreter multivariater Zusammenhänge zwischen Merkmalen innerhalb der Anforderungsprofile für Professorenstellen zu formulieren. Geht man nämlich davon aus, dass es sich bei der Angleichung von Universität und Fachhochschule nur um eine normative Diskursposition handelt, mit der Status, Legitimation und Ressourcenzuflüsse von Fachhochschulen optimiert werden sollen, so müssten sich die faktisch weiterhin bestehenden strukturellen Unterschiede zwischen beiden Hochschultypen in unterschiedlichen Auswahlkriterien bei Berufungen niederschlagen. Ein institutioneller Wandel hätte dieser These zufolge nicht stattgefunden. Vielmehr differieren Uni-

versitäten und Fachhochschulen in vielerlei Hinsicht – und unter anderem eben auch in Bezug auf die Erwartungen, die sie an Kandidaten für eine Professur stellen. Diese Annahme wird im Folgenden als Kern der *Differenzthese* verstanden.

Wird dagegen angenommen, dass es sich bei der Angleichung beider Hochschultypen um ein Realphänomen handelt, das in einer im Gang befindlichen Angleichung der Fachhochschulen an die Universitäten besteht, so ist zu erwarten, dass die Auswahlkriterien für die Besetzung von Professuren an Fachhochschulen denen an Universitäten gleichen – oder doch zumindest stark ähneln – sollten.⁶ Demzufolge hätte ein institutioneller Wandel stattgefunden, der sich unter anderem auf den Bologna-Prozess (Harmonisierung der Abschlüsse) und die durch den angeheizten Wettbewerb um Ressourcen und Legitimation verstärkte Forschungsorientierung der Fachhochschulen zurückführen lässt. Theoretisch ließe sich dieser Prozess mit dem Neo-Institutionalismus als Kombination von erzwungener und mimetischer Isomorphie (*DiMaggio/Powell 1983*) fassen. Dieser institutionelle Wandel (*Walgenbach/Meyer 2008*), der verschiedene Strukturebenen der Fachhochschulen betrifft, müsste sich auch auf der Ebene der Anforderungsprofile für Professorenstellen ausprägen. Diese Annahme macht den Kern der hier als Konvergenzthese bezeichneten Auffassung aus.⁷

Ziel des Artikels ist es vor diesem Hintergrund, einen Beitrag zur Beantwortung der Frage zu leisten, ob die Auswahlkriterien in Berufungsverfahren eher für die Differenz- oder für die Konvergenzthese sprechen. Angestrebt ist nicht, empirische Evidenzen für eine *Entscheidung* zugunsten einer der beiden Thesen an die Hand zu geben. Dafür wären, da die Auswahlkriterien für Professoren nur ein (kleiner) Baustein im Strukturgefüge von Universitäten und Fachhochschulen sind, sehr viel umfangreichere empirische Untersuchungen erforderlich. Auch eine Längsschnittbetrachtung der Entwicklung von Auswahlkriterien im Zeitverlauf ist angesichts des Querschnittsdesigns der Befragung nicht möglich. Und schließlich lässt die Art der Stichprobenziehung keine auf die Erklärung von Kausalzusammenhängen abzielende inferenzstatistische Analyse zu. Anvisiert ist hier vielmehr eine *explorativ ausgerichtete* Analyse, die Indizien dafür liefern soll, ob Universitäten und Fachhochschulen zum Zeitpunkt der Befragung aus Sicht wichtiger Verfahrensbeteiligter eher Gemeinsamkeiten oder Unterschiede bezüglich ihrer kriteriellen Rekrutierungsstandards aufweisen.⁸

⁶Die Konvergenzthese geht dabei realistischerweise nicht von einer bereits vollzogenen Strukturangleichung der Fachhochschulen an die Universitäten aus, sondern nimmt an, dass sich die Unterschiede in Folge der institutionellen Wandlungsprozesse stark abgeschwächt haben.

⁷Der Wissenschaftsrat hatte bereits 2010 „zwei gegenläufige Bewegungen“ identifiziert: „punktuelle Konvergenzen zwischen den Hochschultypen Fachhochschule und Universität und zum anderen zeitlich parallele Ausdifferenzierungsprozesse, und zwar innerhalb der Hochschultypen sowie im Hochschulsystem insgesamt“ (S. 18f.). Die punktuellen Konvergenzen wurden allerdings dahingehend gedeutet, dass sie gerade nicht zu einer Auflösung der Typendifferenz führen.

⁸Dementsprechend wird die multiple Korrespondenzanalyse hier nicht als inferenzstatistische, sondern als explorative Methode eingesetzt.

Ausgehend von den theoretischen Annahmen der Differenz- bzw. Konvergenzthese wird im Folgenden mittels der multiplen Korrespondenzanalyse ein zweidimensionaler Raum der (dichotomisierten⁹) Auswahlkriterien der Berufungskommissionsmitglieder modelliert. In einem weiteren Schritt werden die Ergebnisse der multiplen Korrespondenzanalyse dann mit Hilfe der systemtheoretischen Organisationssoziologie analysiert und im Hinblick auf weitere Forschungsdesiderate diskutiert.

4 Datengrundlage

Grundlage der multiplen Korrespondenzanalyse sind Daten einer Online-Erhebung, die im Rahmen des Projekts „Leistungsbewertung in Berufungsverfahren“ 2015 durchgeführt wurde. Neben Dekanen, Gleichstellungsbeauftragten und Berufungskommissionsmitgliedern wurden dabei auch Berufungskommissionsvorsitzende an allen staatlichen Fachhochschulen und Universitäten (mit Ausnahme von Spezialhochschulen) nach ihrer Einschätzung der Bedeutung von (vorgegebenen) Kriterien in Berufungsverfahren gefragt.

Die Bruttofallzahl dieser Substichprobe beträgt $n=410$ ($n=203$ an Universitäten, $n=207$ an Fachhochschulen). Da Berufungskommissionsvorsitzende nur ein temporäres Amt bekleiden und weder auf Hochschul- noch auf Fakultätsebene öffentlich verzeichnet sind, erfolgte die Stichprobenziehung über eine Kontaktaufnahme zu allen Dekanaten der befragten Hochschulen. Die Dekanate wurden gebeten, den Fragebogen an möglichst alle ihnen bekannten Berufungskommissionsmitglieder weiterzuleiten. Im Fragebogen hatten die Antwortenden dann im Rahmen einer Mehrfachauswahl die Möglichkeit, sich (unter anderem) der Rolle der Berufungskommissionsvorsitzenden zuzuordnen. Allen Respondenten, die diese Möglichkeit genutzt haben, wurde über eine entsprechende Filterführung ein speziell für die Berufungskommissionsvorsitzenden gedachter Fragebogen vorgelegt.

Die Frage zu den Auswahlkriterien für die Kommissionsvorsitzenden bezog sich auf das letzte von ihnen geleitete Verfahren. Sie lautete: „Wie wichtig waren Ihnen persönlich die folgenden Kriterien für die Listenplatzierung der Bewerber?“¹⁰ Die Antwort erfolgte auf einer fünfstufigen Ordinalskala von „nicht wichtig“ bis „sehr wichtig“.¹¹

⁹Die Skalenpunkte 4 und 5 der fünfstufigen Ordinalskala wurden mit „1“ bzw. „+“, die Skalenpunkte 1, 2 und 3 mit „0“ bzw. „-“ codiert.

¹⁰Für jede der befragten Gruppen wurde – bei gleichbleibender Itembatterie – eine andere Frageformulierung gewählt, um den verschiedenen Wissensbeständen und Bewertungsperspektiven der Gruppen gerecht zu werden.

¹¹Auswahl und Formulierung der vorgegebenen Kriterien stützten sich auf eine umfangreiche Sichtung der Forschungsliteratur und des Praxisdiskurses zu Berufungsverfahren, ferner auf 22 Experteninterviews mit verschiedenen an Berufungsverfahren beteiligten Personengruppen und schließlich auf die Ergebnisse eines kognitiven Pretests der Fragebögen. Aufgrund der Gefahr eines Bias infolge von sozialer Erwünschtheit und der damit einhergehenden Verzerrungsgefahr für die Ergebnisse wurden rechtlich illegitime Kriterien (wie Geschlecht – vgl. hierzu instruktiv *Auspurg/Hinz/Schneck 2017* –, Alter, Familienstand, Schicht- (vgl. *Möller 2015*) oder Religionszugehörigkeit, politische Überzeugung, Hautfarbe u.ä.) nicht berücksichtigt.

Bezüglich der Auswahl der Angaben der Berufungskommissionsvorsitzenden für die multiple Korrespondenzanalyse sind zwei Gründe ausschlaggebend. Erstens sind die Vorsitzenden aufgrund ihrer Funktion am umfassendsten über die Verfahren informiert: Sie leiten die Kommissionssitzungen, kommunizieren mit allen am Auswahlverfahren beteiligten Personen und Instanzen und formulieren die Protokolle. Zweitens attestieren sie sich selbst einen im Vergleich mit anderen Akteurgruppen relativ starken Einfluss auf die Listenentscheidung.¹² Aufgrund ihrer Informiertheit und ihres Einflusses kann man daher erwarten, dass ihre Auskünfte zu den Kriterien für die Listenreihenfolge einen prägnanten, wenn auch nicht vollständigen Eindruck der Bewertungsstandards in Berufungsverfahren vermitteln.

5 Korrespondenzanalyse

Im Folgenden wird auf der Basis des skizzierten Datensatzes mithilfe einer multiplen Korrespondenzanalyse geprüft, inwiefern die Relevanzeinschätzung der verschiedenen Kriterien durch die Berufungskommissionsvorsitzenden in einem Zusammenhang mit ihrer Zugehörigkeit zu einem der beiden Hochschultypen steht. Die Aufmerksamkeit gilt also letztlich den durch den Hochschultyp geprägten Antwortmustern der verschiedenen Berufungskommissionsvorsitzenden. Die multiple Korrespondenzanalyse dient dabei primär der Visualisierung und Vereinfachung der komplexen, multivariaten Zusammenhänge zwischen den 22 Indikatoren, die die kriterienbezogenen Präferenzen der Berufungskommissionsvorsitzenden bei Berufungsverfahren abbilden.

Im Rahmen der multiplen Korrespondenzanalyse werden zwei mehrdimensionierte Räume aufgespannt: Zum einen der Raum der Merkmale, zum anderen der Raum der Merkmalsträger (vgl. *Le Roux/Rouanet 2010*). Die multiple Korrespondenzanalyse berücksichtigt dabei sämtliche Zusammenhänge (Interaktionseffekte erster Ordnung) all jener Variablen, die Eingang in die multiple Korrespondenzanalyse gefunden haben (vgl. *Blasius 2010*). Visualisiert wird dabei im Rahmen der durch die multiple Korrespondenzanalyse konstruierten „Raumkarten“, welche Merkmale stark oder weniger stark miteinander assoziiert sind (vgl. *Mundt/Kutzner 2017*). Je häufiger Merkmale gemeinsam zusammen genannt werden, desto stärker bewegen sich die Merkmalspunkte im Korrespondenzraum aufeinander zu und vice versa. Je seltener Merkmalsausprägungen genannt werden, desto mehr rücken sie an die Ränder des Korrespondenzraums (vgl. *Le Roux/Rouanet 2010*). Andersherum gilt: Je häufiger ein Merkmal im Allgemeinen genannt wird, desto näher wandert es in die Nähe des

¹²Laut unserer Online-Erhebung schreiben sich 63 Prozent der Berufungskommissionsvorsitzenden an Universitäten (n=201) und 76 Prozent der Vorsitzenden an Fachhochschulen (n=207) einen großen bis sehr großen Einfluss auf die Listenentscheidung zu.

Zentroids des Korrespondenzraumes, welcher gleichzeitig das Durchschnittsprofil der multiplen Korrespondenzanalyse abbildet (vgl. *Blasius 2001*).¹³

Die multiple Korrespondenzanalyse fungiert hier also als strukturentdeckendes Verfahren, mit dessen Hilfe latente Konstrukte multivariater Daten untersucht werden können. Sie ermöglicht zudem eine Informationsreduktion insbesondere bezüglich der Beschreibung von kategorialen Daten (vgl. *Blasius 1994*). Ziel der multiplen Korrespondenzanalyse ist es dabei, einen optimalen Unterraum eines mehrdimensionierten Hyperraums zu bestimmen, der wiederum ein mögliches Maximum der Variation der Daten erklären kann (vgl. *Blasius 2001*).

Aus Platzgründen muss hier auf eine genauere Darlegung der grundsätzlichen Methodik der multiplen Korrespondenzanalyse verzichtet werden. Für einen kurzen, aber sehr gelungenen und verständlichen Überblick kann die anwendungsorientierte Kurzeinführung von *Mundt/Kutzner (2017)* dienen, für eine ausführlichere Übersicht über die Methode bieten sich die Einführung von *Blasius (2001)* und die etwas aktuellere Darstellung von *Le Roux und Rouanet (2010)* an. Konkrete Anwendungsbeispiele für die multiple Korrespondenzanalyse aus dem Bereich der Hochschulforschung finden sich u. a. bei *Baier/Münch (2013)*, *Bauer/Wolff (2017)*, *Mundt/Kutzner (2017)* und *Schmid/Wilkesmann (2015)*.

5.1 Ergebnisse der Korrespondenzanalyse

Der Raum der Bewertungskriterien in Berufungsverfahren wird durch 22 aktive Variablen mit 44 Ausprägungen (= Anzahl der Spalten der zugrundeliegenden Kreuztabelle¹⁴) und einer Nettotafelzahl von $n=194$ (= Anzahl der Zeilen) aufgespannt. Passiv eingefügt wurde die Variable „Hochschultyp“. Sie wirkt nicht auf die Ausrichtung der Achsen des Korrespondenzraums ein, kann aber gemeinsam mit den „aktiven“ Variablen interpretiert werden (s. u.). Von den $2^{22} = 4.194.304$ möglichen Antwortpattern wurden 176 beobachtet, das häufigste Antwortpattern kommt dabei vier Mal in den Daten vor.

Der Raum der Bewertungskriterien (vgl. Abbildung 1) wird durch zwei Achsen aufgespannt, die grosso modo horizontal und vertikal verlaufen. Die beiden Achsen (bzw. Dimensionen) sind mit einer Varianzaufklärung von 16,10 Prozent und 13,88 Prozent in ihrer Bedeutung für die weitere Analyse in etwa gleichbedeutend. Um die nachfolgenden Analysen für die Leser übersichtlich und nachvollziehbar zu gestalten,

¹³Eine Interpretation der Distanzen im Sinne euklidischer Distanzen ist nicht zulässig (vgl. *Le Roux/Rouanet 2010*, S. 34).

¹⁴Die Kreuztabelle, die hier aus Platzgründen ausgespart bleibt, kann auf Anfrage bereitgestellt werden. Die ursprüngliche fünfstufige Ordinalskala wurde für die multiple Korrespondenzanalyse in eine Dummy-Variablen umcodiert. Die beiden oberen Werte der Ordinalskala („wichtig“ und „sehr wichtig“) werden in der Abbildung 1 durch „+“ (= hohe Relevanz des Kriteriums), die drei unteren Werte durch „-“ (= geringe Relevanz des Kriteriums) wiedergegeben.

wurden die Kategorien der in die multiple Korrespondenzanalyse implementierten Variablen um so dunkler dargestellt, je stärker ihr Beitrag an der Ausrichtung der Achsen des Korrespondenzraumes ausfällt.¹⁵ Auf den Ausschluss von Variablen(-ausprägungen) anhand eines Baseline-Kriteriums (vgl. *Le Roux/Rouanet 2010*) wurde indes verzichtet, nicht zuletzt, um auch jene Merkmale in der Darstellung der multiplen Korrespondenzanalyse zu berücksichtigen, die wenig bis gar keinen Beitrag zu einer der extrahierten Achsen leisten.¹⁶ Die inhaltliche Interpretation der Dimensionen stützt sich im Folgenden jedoch ausschließlich auf jene Indikatoren, die einen überdurchschnittlichen Beitrag zur Ausrichtung der Faktoren leisten (vgl. *Le Roux/Rouanet 2010*).¹⁷

Für die erste Dimension kennzeichnend sind in diesem Zusammenhang auf der linken Seite des Korrespondenzraums deutliche Präferenzen für Merkmale, die dem Aufgabenkomplex der Forschung zugeordnet werden können (vgl. Abbildung 1). Hierzu gehören etwa die Variablen „eingeworbene Drittmittel“, „Mitwirkung in der Scientific Community“, „durchgeführte Forschungsprojekte“, „Betreuung des wissenschaftlichen Nachwuchses“, „Kooperation mit Hochschulen und außeruniversitären Forschungseinrichtungen“ sowie „Forschungskonzept“ und – mit Abstrichen – „wissenschaftliche Publikationen der Bewerber“.¹⁸

Demgegenüber zeichnet sich die erste Achse auf der rechten, äußeren Seite des Korrespondenzraums durch die überdurchschnittlich häufig vorkommende *Absenz* von Präferenzen für die Wichtigkeit von wissenschaftlichen Publikationen, durchgeführten Forschungsprojekten, Forschungskonzept, Kooperation mit Hochschulen und, etwas weniger deutlich, wissenschaftlichem Werdegang, Betreuung wissenschaftlichen Nachwuchses und eingeworbenen Drittmitteln aus.¹⁹ Nimmt man diese Befunde zusammen, so kann man die erste Achse als „Forschungsachse“ bezeichnen, deren Plus-Pol durch Präferenzen für eine hohe Relevanz forschungsbezogener Kriterien und deren Minuspol durch Präferenzen für eine geringe Relevanz forschungsbezogener Kriterien gebildet werden.

¹⁵Im Anhang (vgl. Tabelle 1 bis Tabelle 4) finden sich des Weiteren Auszüge aus den numerischen Ergebnissen der multiplen Korrespondenzanalyse. Die Färbung der 44 verschiedenen Merkmale kann dort indes konkret nachvollzogen werden.

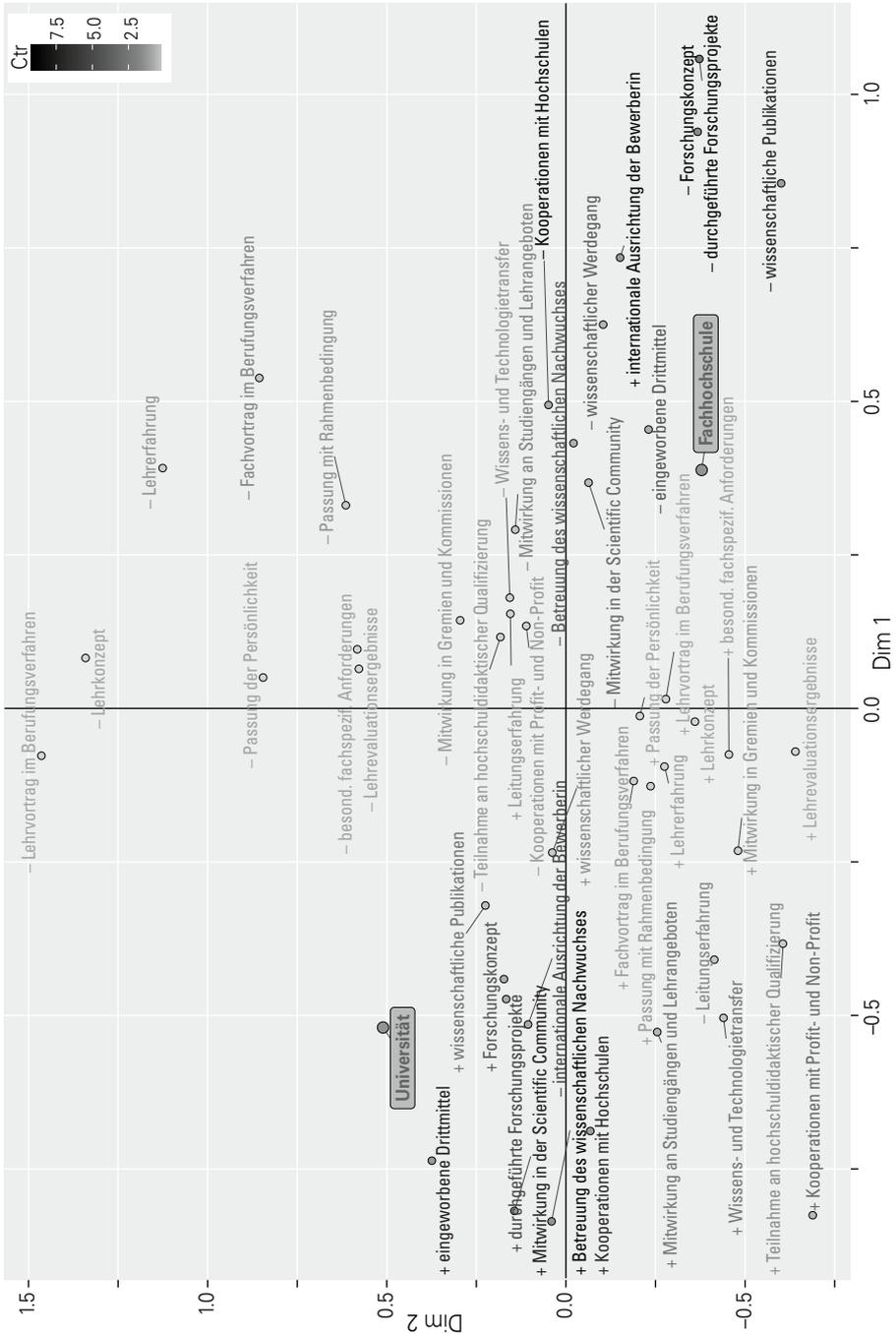
¹⁶Dargestellt werden sollen so z. B. auch jene Indikatoren, die keinen Distinktionswert haben, sich nahe am Zentroid befinden und also allen Berufungskommissionsmitgliedern in etwa „gleich wichtig“ waren.

¹⁷Als „basic criterion“ (vgl. *Le Roux/ Rouanet 2010*, S. 99) kann auf beiden Dimensionen ein Beitrag von $100/46 = 2,17$ Prozent geltend gemacht werden.

¹⁸Über Gründe dafür, dass das Kriterium „Mitwirkung an Studiengängen und Lehrangeboten“ leicht überdurchschnittlich zur Achse der Forschung beiträgt und zudem deutlicher mit der Universität als der Fachhochschule assoziiert ist, lässt sich nur – zum Beispiel mit Blick auf die größere Selbstverständlichkeit der Mitwirkung an der Lehre an Fachhochschulen, die nicht eigens hervorgehoben werden muss – spekulieren.

¹⁹Einen ähnlich starken Beitrag wie die vorgenannten Variablen zur Konfiguration der 1. Dimension weist die Variable „internationale Ausrichtung der Bewerberin“ auf. Diese ist inhaltlich nicht eindeutig der Gruppe der forschungsbezogenen Variablen zuzurechnen. Eine mögliche Erklärung für diese Konfiguration könnte darin liegen, dass die Fachhochschulen im Bereich der Internationalisierung einen Nachholbedarf sehen und z. B. das Abhalten englischsprachiger Veranstaltungen explizit als Anforderung in den Ausschreibungstexten nennen.

Abbildung 1: Raum der Bewertungskriterien – Punktwolke der 22 x 2 = 44 Ausprägungen (grau abgestuft: Beitragsstärke der Merkmale zur 1. Dimension)

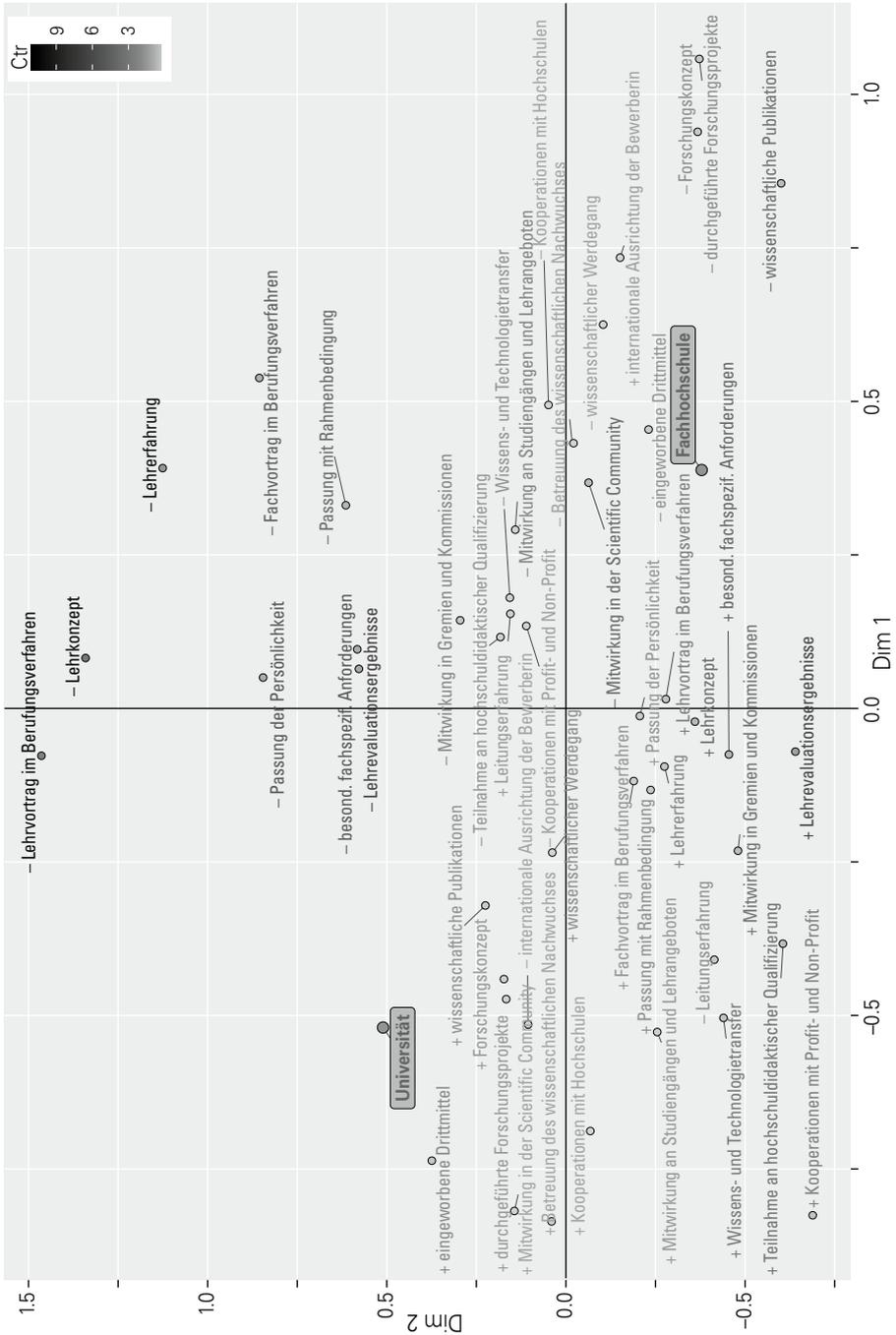


Die zweite Achse des Korrespondenzraums konstituiert sich primär anhand der Differenzen zwischen Variablen, die der Dimension „Lehre“ zugeordnet werden können (vgl. Abbildung 2). So findet sich im unteren Teil des Korrespondenzraums eine deutliche Präferenz für eine hohe Relevanz der Kriterien „Lehrevaluationsergebnisse“ und „Lehrkonzept“. Demgegenüber zeichnet sich in der oberen Peripherie des Korrespondenzraums eine Dominanz von Präferenzen ab, die die geringe Bedeutung von Lehrkriterien für die Listenentscheidung betonen. Dies betrifft vor allem die Variablen „Lehrvortrag im Berufungsverfahren“, „Lehrkonzept“, „Lehrevaluationsergebnisse“ und „Lehrerfahrung“. Auch wenn nicht alle die Lehre betreffenden Variablen gleichermaßen stark zur Achsenausrichtung beitragen, kann das untere Achsenende der zweiten Dimension des Korrespondenzraums als Pluspol, das obere Achsenende hingegen als Negativpol der Dimension „Relevanz der Lehre“ verstanden werden.

Um nun die Zugehörigkeit der Berufungskommissionsvorsitzenden zu einem der beiden Hochschultypen im Hinblick auf die Konfiguration des Korrespondenzraums berücksichtigen zu können, wurde die Variable „Hochschultyp“ als passive, illustrierende Variable in den Korrespondenzraum eingefügt. Die multiple Korrespondenzanalyse erlaubt es, beliebige Indikatoren als „passive“ Variablen in die Analyse einzubeziehen. Der Vorteil letzterer besteht darin, dass sie die geometrische Ausrichtung der Achsen der multiplen Korrespondenzanalyse nicht beeinflussen (vgl. *Blasius 2001*), wohl „[...] aber gemeinsam mit den anderen Merkmalen interpretiert werden [können]“ (*ibd.*, S.85) können. „Illustrierende“ bzw. „passive“ Variablen (*Blasius 2001*, S. 144) sind nun, verstanden als „strukturierende Faktoren“, imstande, die Konfiguration des durch die multiple Korrespondenzanalyse aufgespannten Raums zu „erklären“. Auch wenn im Rahmen der multiplen Korrespondenzanalyse das Begriffspaar „abhängige/unabhängige Variable“ nicht zulässig ist, lässt sich die Konfiguration des Korrespondenzraums in gewisser Weise dennoch als abhängige, der strukturierende Faktor als unabhängige Variable verstehen (vgl. *Le Roux/Rouanet 2010*).

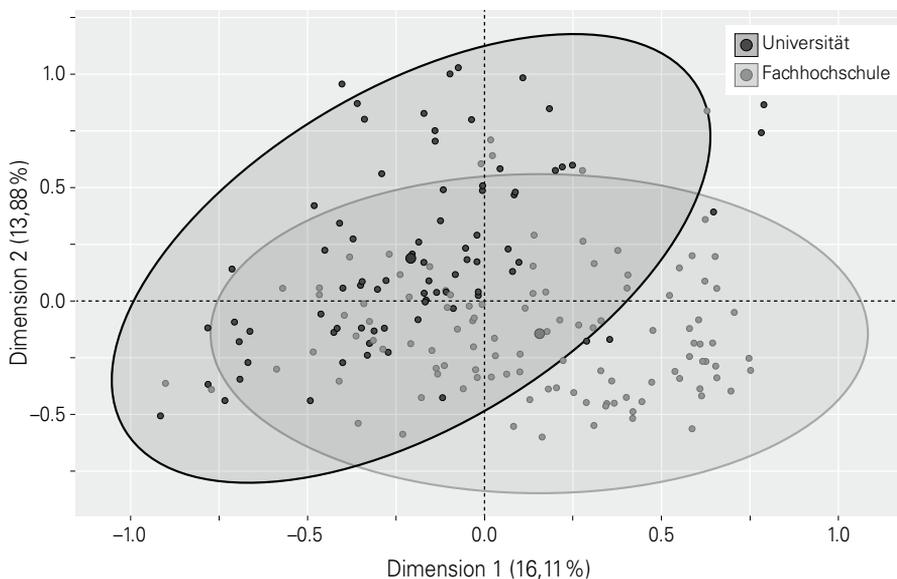
Bezieht man nun die Hochschulzugehörigkeit der Berufungskommissionsvorsitzenden als strukturierenden Faktor in die multiple Korrespondenzanalyse ein, so zeigt sich eine recht deutliche Separation zwischen den Mitgliedern der Fachhochschulen und denen der Universität (Abbildung 1 o. 2). Deutlich wird, dass die Berufungskommissionsvorsitzenden der Universitäten dem Pluspol der Dimension Forschung einerseits, dem Minuspol der Dimension Lehre andererseits zugeneigt sind. Den Berufungskommissionsvorsitzenden der Universitäten diametral gegenüber situiert, zeichnen sich die Berufungskommissionsvorsitzenden der Fachhochschulen durch eine ausgeprägte Neigung zum Pluspol der Dimension Lehre bei gleichzeitiger Nähe zum Minuspol der Dimension Forschung aus. Die Separierung der Zugehörigkeit zu einem der beiden Hochschultypen über die Dimensionen Lehre und Forschung hinweg ist mit einem η^2 -Koeffizienten von je 0,20 über beide extrahierten Dimensionen hinweg gleich stark und könnte pointiert unter der Überschrift „Forschung vs. Lehre“ rubriziert werden.

Abbildung 2: Raum der Bewertungskriterien – Punktwolke der 22 x 2 = 44 Ausprägungen (grau abgestuft: Beitragsstärke der Merkmale zur 2. Dimension)



Eine perfekte Separierung, die sich darin äußern würde, dass sich die Konzentrationsellipsen der Fachhochschul- und Universitäts-Berufungskommissionsvorsitzenden nicht überschneiden, liegt indes nicht vor. Dieser Umstand ist nicht zuletzt darauf zurückzuführen, dass sich bei drei der weiteren Kriterien keine eindeutige Assoziation mit dem Hochschultyp ausmachen lässt. Dies gilt für die Kriterien „Wissens- und Technologietransfer“, „Kooperation mit Profit-/Nonprofit-Organisationen“ und „Mitwirkung in Gremien und Kommissionen der akademischen Selbstverwaltung“, deren Lokalisation im unteren linken und oberen rechten Quadranten in der ersten Dimension der Lokalisation der Universität, der zweiten Dimension dagegen derjenigen der Fachhochschule nahekommt (vgl. Abb. 3).

Abbildung 3: Raum der Bewertungskriterien – Punktwolke der 194 Berufungskommissionsvorsitzenden (Konzentrationsellipsen der Berufungskommissionsvorsitzenden der Universitäten und der Fachhochschulen)



Bei drei weiteren Kriterien sprechen die Ergebnisse der multiplen Korrespondenzanalyse dafür, dass die Annahmen der Differenzthese eher nicht zutreffen. So werden die besonderen fachspezifischen Anforderungen nicht in gleicher Weise von Universitäten und Fachhochschulen gewichtet, sondern ihre hohe Relevanz wird überdurchschnittlich häufig von Fachhochschul-Vorsitzenden, ihre geringer Relevanz von Vorsitzenden von Universitäten betont. Ein Grund dafür könnte die besondere Problematik des Nachweises außerhochschulischer Praxiserfahrungen als *notwendiges* Kriterium bei der Berufung von Fachhochschulprofessoren sein (*In der Smitten 2017, S. 26*). Analoges gilt für das Kriterium der Leitungserfahrungen, auch hier messen die Fachhochschulen

dem Kriterium überdurchschnittlich häufig eine große Bedeutung zu. Die internationale Ausrichtung wird entgegen der Erwartung der Differenzthese überdurchschnittlich häufig von den Fachhochschulmitgliedern als wichtig eingeschätzt – und vice versa. Ob dies – wie oben in der Fußnote 19 vermutet – auf einen besonderen Nachholbedarf der Fachhochschulen in diesem Bereich zurückzuführen ist, muss hier offen bleiben.

Für die Differenzthese sprechen schließlich die zwei letzten Kriterien: So trägt das Kriterium der „Kooperation mit Hochschulen und außerhochschulischen Forschungseinrichtungen“ stark zur ersten Dimension („Forschung“) bei und ist in seiner Einschätzung als „wichtig“ deutlich mit der Universität assoziiert, in seiner Einschätzung als „nicht wichtig“ hingegen mit der Fachhochschule. Die Passung der Persönlichkeit der Bewerber zum Kollegium dagegen, die überdurchschnittlich zur Dimension „Lehre“ beiträgt, wird überdurchschnittlich häufig von den Berufungskommissionsvorsitzenden der Fachhochschulen als relevantes Auswahlkriterium hervorgehoben.

5.2 Interpretation der Befunde aus organisationssoziologischer Perspektive

Um nun die Befunde mit Blick auf die Differenz- und die Konvergenzthese bewerten und theoretisch einordnen zu können, werden sie im Folgenden aus der Perspektive der systemtheoretischen Organisationssoziologie analysiert. Dabei ist der Grundgedanke, dass sich die Organisationsstrukturen von Hochschulen – d. h. insbesondere ihre Aufgaben bzw. „Zweckprogramme“ (siehe unten) – in den Anforderungsprofilen von Professuren widerspiegeln. Angenommen wird zudem, dass die Berufungskommissionsvorsitzenden über die Anforderungsprofile für die Professuren und die daraus abzuleitenden Konsequenzen plausibel Auskunft geben können, weil sie über das entsprechende Betriebswissen verfügen und den auf der Basis der Stellenausschreibung erfolgenden Auswahlprozess kontinuierlich begleiten. Welche Ausprägungen von Organisationsstrukturen der beiden Hochschultypen schlagen sich nun in den Kriterien bei der Auswahl für eine Professur an Fachhochschulen bzw. Universitäten nieder?

Wenn man mit der Systemtheorie davon ausgeht, dass Hochschulen formale Organisationen sind, die sich über Entscheidungen reproduzieren und deren Strukturen „Entscheidungsprämissen“ (Simon 1981 [1945]) darstellen, so lassen sich mit Luhmann (2006) und Kühl (2011) drei formale Strukturkomponenten unterscheiden: *Entscheidungsprogramme* legen Rahmenbedingungen für Art und Richtigkeit des Entscheidens in der Organisation fest. Dabei wird zwischen *Zweckprogrammen* (Aufgaben und Ziele der Organisation) und *Konditionalprogrammen* (Wenn-Dann-Regeln für die Kopplung spezifischer Inputs an bestimmte Verarbeitungsprozeduren) unterschieden. Im Folgenden werden aufgrund ihrer Relevanz für die Auswahlkriterien nur die Zweckprogramme beider Hochschultypen, d. h. insbesondere Forschung und Lehre betrachtet.

Kommunikationswege als eine weitere formale Strukturkomponente definieren Weisungsrechte und Berichtspflichten in der Organisation. Sie legen zugleich die Außengrenzen der Organisation (die mit den Grenzen des Entscheidens über die Mitglieder zusammenfallen) wie auch ihre interne vertikale und horizontale Gliederung (z. B. Fakultäten Entscheidungsebenen) fest. Die Mitarbeiter schließlich (Strukturkomponente „*Personal*“) prägen durch ihre Eigenschaften und Verhaltensmuster die Art des Entscheidens in der Organisation mit.²⁰

Vor diesem Hintergrund kann die Differenzthese so verstanden werden, dass sie die *Unterschiede* zwischen den Strukturkomponenten von Universitäten und Fachhochschulen sowie die daraus ableitbaren Unterschiede zwischen den jeweiligen Auswahlkriterien betont. Dagegen unterstreicht die Konvergenzthese in organisationssoziologischer Lesart die durch den institutionellen Assimilationsprozess erzeugten *Strukturähnlichkeiten* zwischen beiden Hochschularten, die eher für Gemeinsamkeiten zwischen den Auswahlkriterien sprechen.

Wie lassen sich nun die in Abschnitt 4 im Rahmen der multiplen Korrespondenzanalyse herausgearbeiteten Befunde im Hinblick auf die Strukturähnlichkeiten bzw. -differenzen der beiden Hochschultypen verstehen? Die Befunde bezüglich der beiden herausgearbeiteten zentralen Dimensionen bzw. Achsen „Forschung“ und „Lehre“ und bezüglich der Assoziation letzterer mit den passiv hinzugefügten Merkmalen „Universität“ bzw. „Fachhochschule“ lassen den Schluss zu, dass sich die bisherige institutionelle „Arbeitsteilung“ zwischen Universitäten und Fachhochschulen auf der Ebene ihrer Berücksichtigung in den Kriterien für die Besetzung von Professuren noch nicht entscheidend gewandelt hat. Dass die Fachhochschulen auf der Ebene der obersten Zweckprogramme nach wie vor am Primat der Lehre, die Universitäten hingegen am Primat der Forschung orientiert sind, tritt – wenn man die Homologie von organisationalem und stellenbezogenem Aufgabenprofil als Annahme gelten lässt – in der multiple Korrespondenzanalyse sehr deutlich heraus. Dabei wird durch die starke Assoziation der Universitäten mit der Forschung auch offenkundig, dass alle Bekenntnisse zu einer hohen Relevanz der Lehre für Universitäten und mithin für die Profile von Kandidaten für eine Universitätsprofessur bislang eher Lippenbekenntnisse sind. Offenbar werden lehrbezogene Kriterien zwar auch an Universitäten in Anschlag gebracht, gegenüber forschungsbezogenen Kriterien aber deutlich geringer gewichtet. Die Befunde lassen damit den Schluss zu, dass bezüglich der Zweckprogrammatisierung der professoralen Stelle (ihres Aufgabenprofils) die Differenzthese zutrifft: So unterscheiden sich die aufgabenbezogenen Anforderungen an die Kandidaten für Fachhoch-

²⁰Ein alternatives, aber ähnlich gelagertes organisationales Strukturmodell entwirft *Würmseer (2010)* für die Analyse der Identität von Fachhochschulen und Universitäten.

schulprofessuren erheblich von denen, die an Kandidaten für Universitätsprofessuren gerichtet werden. Hier schlägt die bisherige, tradierte Dominanz der Forschung an Universitäten und der Lehre an Fachhochschulen durch (*Würmseer 2010; Vogel 2009*), eine Ungleichverteilung, die sich strukturell vor allem über die unterschiedlichen Lehrdeputate (18 SWS an Fachhochschulen,²¹ 8 bis 9 SWS an Universitäten)²², lehrbezogenen Personalstrukturen (*Bloch et al. 2014*), das eingeschränkte Fächerspektrum der Fachhochschulen²³ und das noch weitgehend den Universitäten vorbehalten Promotions- (und Habilitations-)Recht reproduziert und die sich anhand der Unterschiede zwischen beiden Hochschultypen bei den Forschungsausgaben²⁴ oder bei den Drittmittelquellen²⁵ auch empirisch ablesen lässt.

Allerdings erlauben die Befunde und die Datengrundlagen der multiplen Korrespondenzanalyse keine Längsschnittbetrachtung, mit der sich mögliche Veränderungen von professoralen Aufgabenprofilen und den entsprechenden Kriterien in Berufungsverfahren abbilden ließen. Es kann daher nicht ausgeschlossen werden, dass die Assoziation von forschungsorientierten Auswahlkriterien mit der Universität und von lehrbezogenen Auswahlkriterien mit der Fachhochschule bei einer Untersuchung zu einem Zeitpunkt in der jüngeren Vergangenheit noch stärker ausgefallen wäre. Für eine solche Vermutung spricht – auf der Ebene des Fachhochschulsystems – die (umstrittene)²⁶ Übertragung des Promotionsrechts auf forschungsstarke Einrichtungen an den Fachhochschulen in Hessen und auch die Entwicklung bei quantitativen, forschungsassoziierten Indikatoren wie die deutlich ansteigende Quote der Fachhochschulausgaben für Forschung und Entwicklung²⁷ und der Drittmittel²⁸ sowie die kräftig steigende Förderung der Forschung an Fachhochschulen durch das Bundesministerium

²¹Eine Ausnahme ist Sachsen-Anhalt mit 16 SWS.

²²Zum (zu knappen) Zeitbudget für die Forschung siehe: HLB o.J., S. 16.

²³Die nach 1968 gegründeten Fachhochschulen sind vor allem aus ingenieurwissenschaftlichen Bildungseinrichtungen hervorgegangen und haben später die Wirtschafts- und die angewandten Sozialwissenschaften an sich gezogen. Eine weitere Ausdifferenzierung des Fächerspektrums ist – von kleinen Ausnahmen abgesehen – ausgeblieben (vgl. *Schreiterer 2016*). Dagegen decken Universitäten das gesamte disziplinäre Fächerspektrum der Wissenschaft ab.

²⁴So lagen die Ausgaben für Forschung und Entwicklung im Jahr 2014 bei den Universitäten bei 10,1 Mrd. Euro, bei den Fachhochschulen dagegen bei 853,6 Mio. Euro (*Statistisches Bundesamt 2016a*, S. 375). Vgl. auch *Schiller/Mahmud/Kenkel 2015*, S. 22.

²⁵Größter Drittmittelgeber im Jahr 2011 war bei den Fachhochschulen der Bund (40,9% der Drittmittel), bei den Universitäten die DFG (36,3% der Drittmittel). Vgl. *Hachmeister/Duong/Roessler 2015*, S. 10.

²⁶Vgl. *Hartmer/Detmer S. 245f.*

²⁷So sind die Ausgaben der Fachhochschulen (inkl. Verwaltungsfachhochschulen) von 645,266 Mio. Euro im Jahr 2011 auf 853,632 Mio Euro im Jahr 2014 angestiegen (*Statistisches Bundesamt 2016a*, S. 375).

²⁸Die Drittmittel der Fachhochschulen sind von 299 Mio. (2009) auf 454 Mio. (2012) gestiegen, die Drittmittel pro Professor im gleiche Zeitraum von 20.400 auf 27.100 Euro (*Hachmeister/Duong/Roessler 2015*, S. 5f.).

für Bildung und Forschung (BMBF).²⁹ Dennoch fallen bei der Querschnittsbetrachtung im Rahmen der multiplen Korrespondenzanalyse die Unterschiede bei den forschungs- und lehrbezogenen Kriterien zwischen beiden Hochschultypen so deutlich aus, dass nach wie vor von einer klaren Differenz der Aufgabenprofile auf der Ebene der Hochschule wie der einzelnen Professuren auszugehen ist.

Was die Kommunikationswege betrifft (das heißt die Außengrenzen der Hochschulen sowie ihre vertikale und horizontale Binnengliederung), fällt in Bezug auf das Kriterium der Leitungserfahrung auf, dass ihm an Fachhochschulen überdurchschnittlich häufig eine große Bedeutung zuerkannt wird. Dies könnte zwar in Anbetracht des Umstands, dass Fachhochschulprofessuren nur in seltenen Fällen mit Personalführungsaufgaben bezüglich wissenschaftlicher Mitarbeiterinnen und Mitarbeiter betraut sind, überraschen. Da Leitungserfahrung im Fragebogen allerdings im Sinne von Erfahrungen als „Studiendekan(in), als Leiter(in) eines Sonderforschungsbereichs oder als Abteilungsleiter(in) in einem Unternehmen“ erläutert worden war, lässt sich das Ergebnis erklären. Denn wenn man sich vor Augen führt, dass die Professorenschaft an Fachhochschulen aufgrund des deutlich höheren Koordinationsbedarfs zwischen den Professuren in der Lehre und aufgrund der weitgehenden Absenz eines wissenschaftlichen Mittelbaus eher wie das Kollegium in einer weiterführenden Schule strukturiert ist (und nicht die aus Universitäten bekannte Struktur einer losen Kopplung je für sich operierender „Lehrstühle“ aufweist) berücksichtigt, kann man die Bedeutung des Kriteriums Leitungserfahrung auf die besondere Notwendigkeit der Übernahme herausgehobener Positionen in der akademischen Selbstverwaltungshierarchie – insbesondere auch als „Studiendekan“ – zurückführen. Während an Universitäten jede Professorin bzw. jeder Professor (fach- und ausstattungsabhängig) in unterschiedlichem Maße Untergebene, nämlich die wissenschaftlich Mitarbeitenden, führt und die Anforderung nach Personalführungskompetenz daher gar nicht mehr eigens mitbedacht werden muss, besteht die besondere Führungsherausforderung an Fachhochschulen darin, die professoralen Kolleginnen und Kollegen im Rahmen der Produktion von nur kollektiv realisierbaren Leistungen (Abschluss eines Studiengangs) zu führen.³⁰ Dies könnte sich an dieser Stelle in den besonderen Anforderungen an die Leitungserfahrung niederschlagen. Hinzu kommt die besondere Bedeutung, die außerhochschulische Berufs- und Leitungserfahrungen für die Fachhochschulkarriere besitzen; sie werden durch das Beispiel der „Abteilungsleiter in einem Unternehmen“ angesprochen und weiter unten im Zusammenhang mit der Strukturkomponente des Personals separat behandelt.

²⁹Das Fördervolumen des BMBF-Programms „Forschung an Fachhochschulen“ hat sich von 10,5 Mio. Euro im Jahr 2005 (Haushaltsansatz) auf 55 Mio. Euro im Jahr 2017 mehr als verfünffacht. Vgl. <https://www.bmbf.de/de/forschung-an-fachhochschulen-543.html>, zuletzt abgerufen am 25.9.2017. Hierzu und zu weiteren Förderprogrammen Borgwardt 2016, S. 51ff.

³⁰2016 lag die Zahl der hauptberuflichen wissenschaftlichen und künstlerischen Mitarbeiter an Universitäten bei 167.952, an Fachhochschulen (ohne Verwaltungsfachhochschulen) dagegen bei 11.865 (*Statistisches Bundesamt 2016b, S. 92*).

Auf die gegenüber den Universitäten anders gelagerte Personalstruktur der Fachhochschulen lässt sich ggf. auch zurückführen, dass die Passung der Persönlichkeit der Bewerberinnen und Bewerber zum Kollegium überdurchschnittlich zur Dimension „Lehre“ beiträgt und überdurchschnittlich häufig von den Berufungskommissionsvorsitzenden der Fachhochschulen als relevantes Auswahlkriterium hervorgehoben wird. Es ist gerade die an Fachhochschulen unverzichtbare kollegiale Kooperation in der Lehre (hier verstanden als Spezifikum der Kommunikationswege an den Fachhochschulen), die es – zusammen mit der insgesamt geringeren Größe der Fachhochschulen³¹ und der dadurch bedingten Formen der Interaktion und Entscheidungskommunikation – besonders erforderlich macht, auf die Passung der Person zum Kollegium zu achten. Auf Quertreiber, Drückeberger oder Eigenbrötler hat man hier noch weniger gewartet als an Universitäten, wo die oftmals vollständige wechselseitige Indifferenz der Professuren in der Forschung einen größeren Spielraum für persönliche Differenzen lässt.

Was schließlich die Außengrenzen der hochschulischen Kommunikationswege angeht, zeigt die multiple Korrespondenzanalyse, dass das Kriterium der „Kooperation mit Hochschulen und außerhochschulischen Forschungseinrichtungen“ stark zur ersten Dimension („Forschung“) beiträgt und in seiner Einschätzung als „wichtig“ deutlich mit der Universität assoziiert ist, während sich seine Einschätzung als „nicht wichtig“ stärker mit der Fachhochschule verbindet. Zum einen bestätigt sich hier noch einmal der Unterschied in den Aufgabenprofilen beider Hochschultypen: Während die Universitäten die Forschung priorisieren (und entsprechend ihr professorales Personal rekrutieren), steht bei den Fachhochschulen (nach wie vor) die Lehre im Mittelpunkt. Diese Differenz bei den Zweckprogrammen bedingt dann auch den Unterschied bezüglich der Relevanz der Kooperationsbeziehungen zu anderen Hochschulen und außeruniversitären Forschungseinrichtungen, die man einerseits bei Bewerbungen für Universitätsprofessuren und andererseits bei Bewerbungen für Fachhochschulprofessuren erwartet.

Fasst man die Überlegungen zu den Kriterien zusammen, die sich (schwerpunktmäßig) auf die Kommunikationswege von Universitäten und Fachhochschulen beziehen, so zeigt sich auch hier, dass die Ergebnisse der multiplen Korrespondenzanalyse eher *für die Differenzthese* sprechen. Der vergleichsweise hohe Stellenwert von „Leitungserfahrung“ und von „Passung zum Kollegium“ wie auch der vergleichsweise geringe Stellenwert von „Kooperationserfahrungen mit anderen Hochschulen und außeruniversitären Forschungseinrichtungen“ bei Fachhochschulen lässt sich auf die Differenzen in den Kommunikationswegen zurückführen, die zwischen Fachhochschulen und Universitäten bestehen.

³¹ Staatliche Universitäten ohne Spezialhochschulen haben im Mittel rund 20.000 Studierende, Fachhochschulen (gleichfalls ohne Spezialhochschulen) dagegen rund 6.000.

Einen Hinweis darauf, dass die Differenzthese zutrifft, liefern schließlich auch die Befunde der multiplen Korrespondenzanalysen, die man der Strukturkomponente des Personals zurechnen kann. Dabei geht es hier nicht um ein besonderes Persönlichkeitsprofil, das man den Bewerberinnen und Bewerbern an Universitäten bzw. an Fachhochschulen abverlangte, sondern um Spezifika der *beruflichen Vorerfahrungen*, die diese mitbringen müssen, um an einer Fachhochschule berufbar zu sein. Diese Spezifika für Fachhochschulprofessoren sind in den Landeshochschulgesetzen formal festgeschrieben: So fordert z.B. § 25 des Niedersächsischen Hochschulgesetzes „besondere Leistungen bei der Anwendung oder Entwicklung wissenschaftlicher Erkenntnisse und Methoden in einer mindestens fünfjährigen beruflichen Praxis, von der mindestens drei Jahre außerhalb des Hochschulbereichs ausgeübt worden sein müssen.“ Aus empirischen Untersuchungen (*In der Smitten, Sembritzki, Thiele 2017a* und *2017b; In der Smitten et. al. 2017*) ist bekannt, dass die Erfüllung dieses notwendigen Kriteriums bei der Besetzung von Fachhochschulprofessuren oft Schwierigkeiten bereitet. Dies ist nicht zuletzt darauf zurückzuführen, dass die Fachhochschulen die Erzeugung ihres eigenen professoralen Nachwuchses nicht autonom bewerkstelligen können, sondern diesbezüglich auf die Universitäten (Promotion) und auf wissenschaftsexterne Einrichtungen (Berufspraxis) angewiesen sind (*In der Smitten et al. 2017; Vogel 2009*).

In diesem Zusammenhang zeigt die multiple Korrespondenzanalyse, dass sich diese Schwierigkeiten der Reproduktion des Nachwuchses auf die Gewichtung des Kriteriums der besonderen fachspezifischen Karrierevoraussetzungen an Fachhochschulen und Universitäten auswirken. Obwohl auch Universitäten besondere fachspezifische Anforderungen haben (z. B. Erfahrungen in der Krankenversorgung im Bereich medizinischer Professuren),³² werden die besonderen fachspezifischen Anforderungen an Fachhochschulen stärker gewichtet als an Universitäten. So wird ihre hohe Relevanz überdurchschnittlich häufig von Berufungskommissionsvorsitzenden an Fachhochschulen geltend gemacht, während die Vorsitzenden von Berufungskommissionen an Universitäten überdurchschnittlich häufig ihre geringe Bedeutsamkeit betonen. In diesen divergierenden Anforderungen spiegeln sich die unterschiedlichen Karrierewege beider Einrichtungstypen wider, die zur organisationalen Strukturkomponente des Personals gehören. Folglich zeichnet sich auch in Bezug auf diese Strukturkomponente ab, dass eine organisationssoziologische Lesart der Befunde der multiplen Korrespondenzanalyse eher Argumente für die Differenzthese als für die Konvergenzthese an die Hand gibt.

³²Das Auswahlkriterium „besondere fachspezifische Anforderungen“ wurde im Fragebogen sowohl im Hinblick auf universitätstypische Anforderungen („Erfahrungen in der Krankenversorgung“) als auch auf Anforderungen, die für Fachhochschulen einschlägig sind („sonstige berufspraktische Erfahrungen“), erläutert.

6 Fazit und Diskussion

Die Ergebnisse der multiplen Korrespondenzanalyse zeigen, dass die Berufungskommissionsvorsitzenden von Universitäten im Vergleich zu ihren Fachhochschulkollegen die Relevanz forschungsbezogener Kriterien (wissenschaftliche Publikationen, durchgeführte Forschungsprojekte, eingeworbene Drittmittel, die Mitwirkung in der Scientific Community (über die Tätigkeit als Vorstand einer Fachgesellschaft oder als Herausgeber von Fachzeitschriften) sowie das Forschungskonzept bzw. geplante Forschungsprojekte und der Fachvortrag im Berufungsverfahren) für die Listenentscheidung in Berufungsverfahren überdurchschnittlich häufig betonen. Auf der anderen Seite akzentuieren sie überproportional häufig die geringe Relevanz von lehrbezogenen Kriterien (Lehrerfahrung bezüglich Veranstaltungen, Prüfungen und der Betreuung von Abschlussarbeiten; Lehrevaluationsergebnisse; das Lehrkonzept; schließlich der Lehrvortrag im Berufungsverfahren – zweite Dimension). Für die Berufungskommissionsvorsitzenden von Fachhochschulen zeigt sich dagegen ein inverses Bild: Sie heben überdurchschnittlich häufig hervor, dass forschungsbezogene Kriterien bei der Listenentscheidung keine besondere Rolle spielen, und betonen dagegen die Relevanz lehrbezogener Kriterien. Auf der Ebene der Zweckprogramme der Hochschule wie der Professur zeigt sich also eine deutliche Dominanz der Forschung bei den universitären Verfahren, während bei den Fachhochschulen (nach wie vor) die Lehre im Vordergrund steht.

Bezüglich der Kommunikationswege von Universitäten und Fachhochschulen wurde vor dem Hintergrund der empirischen Befunde argumentiert, dass sich die Unterschiede bei der Gewichtung bestimmter Auswahlkriterien auf die strukturellen Unterschiede bei den Kommunikationswegen beider Einrichtungstypen zurückführen lassen. Dies betrifft den hohen Stellenwert der Kriterien „Leitungserfahrung“ und „Passung zum Kollegium“ wie auch den vergleichsweise geringen Stellenwert des Kriteriums „Kooperationserfahrungen mit anderen Hochschulen und außeruniversitären Forschungseinrichtungen“ in Fachhochschul-Berufungsverfahren.

Auf der Ebene des Personals schließlich werden Unterschiede zwischen Fachhochschulen und Universitäten bei der Gewichtung der besonderen beruflichen Vorerfahrungen der Bewerber deutlich, die gesetzlich zwingende Eintrittsbedingungen darstellen. Hier dürften es die spezifischen Anforderungen im Rahmen einer Fachhochschulprofessur-Karriere sein (nämlich die fünfjährige berufliche und – davon – mindestens dreijährige außer-hochschulische Berufserfahrung), die für die Gewichtungsunterschiede verantwortlich sind.

Aus diesen Befunden lässt sich nun eine Antwort auf die oben aufgeworfene Frage ableiten, inwiefern die Angaben der Berufungskommissionsvorsitzenden zur Relevanz

von Kriterien bei der Listenentscheidung in Berufungsverfahren eher für die Differenz- oder für die Konvergenzthese sprechen. Offenkundig stützen die Ergebnisse der multiplen Korrespondenzanalyse die Annahme, dass nach wie vor erhebliche strukturelle Unterschiede zwischen Fachhochschulen und Universitäten bestehen, die sich in Unterschieden bei den Auswahlkriterien manifestieren. So gewichten Berufungskommissionsvorsitzende an Fachhochschulen lehrbezogene Kriterien und Berufungskommissionsvorsitzende von Universitäten forschungsbezogene Kriterien bei der Listenplatzierung deutlich stärker. Dies entspricht den traditionell divergierenden Aufgabenprofilen der beiden Hochschultypen, denen gemäß die Universität primär für die Forschung, die Fachhochschule primär für die Lehre zuständig ist. Ähnliche Übereinstimmungen lassen sich zwischen den Kommunikationswegen und Karrierestrukturen beider Einrichtungstypen und einzelnen der abgefragten Auswahlkriterien identifizieren. Dementsprechend sprechen die Analyseergebnisse recht klar für die Differenzthese. Ein Wandel der institutionalisierten Aufgabenverteilung zwischen Universitäten und Fachhochschulen lässt sich auf der Basis der Analyse der Befragungsergebnisse der Kommissionsvorsitzenden aus dem Jahr 2015 jedenfalls nicht erkennen.

Wo liegen die Grenzen dieser Aussage? Vier Limitationen (neben der fehlenden inferenzstatistischen Prüfung) sind hier zu nennen. Erstens wird mit den Kriterien für die Listenentscheidung in Berufungsverfahren nur ein spezifischer Merkmalskomplex von Fachhochschulen und Universitäten betrachtet. Zwar ist die Rekrutierung von Professoren eine für die Hochschulen strategisch wichtige Aufgabe. Für eine Entscheidung zwischen der Differenz- und der Konvergenzthese müssten aber letztlich alle wesentlichen Strukturmerkmale der beiden Hochschultypen verglichen werden.

Zweitens bezieht sich die Analyse nur auf die Angaben von Berufungskommissionsvorsitzenden. Auch für die Ebene der Auswahlkriterien bleibt das Bild daher fragmentarisch. Zwar lässt sich für die Berufungskommissionsvorsitzenden annehmen, dass ihre Angaben zu den Auswahlkriterien aus den oben genannten Gründen besonders aussagekräftig sind. Dennoch wäre es für ein vollständiges Bild erforderlich, auch die übrigen am Auswahlprozess beteiligten Personengruppen in die Analyse mit einzubeziehen. Dies war im vorliegenden Projektkontext aufgrund der unterschiedlichen Frage-Formulierungen für Berufungskommissionsvorsitzende, Berufungskommissionsmitglieder, Gleichstellungsbeauftragte und Dekane methodisch nicht möglich.

Drittens besteht ein grundsätzliches methodologisches Problem der Analyse von Auswahlkriterien bei Personalentscheidungen darin, dass die Anwendung von Auswahlkriterien nicht direkt beobachtbar ist. Man ist entweder auf Selbstauskünfte der Verfahrensbeteiligten (wie im vorliegenden Fall; vgl. *Abele-Brehm, Bühner 2016*) oder aber auf die Interpretation ihrer Kommunikationen und Handlungen angewiesen. Beide

Forschungsstrategien sind nicht unproblematisch: bei Selbstauskünften besteht die Gefahr einer bewussten oder unbewussten Verzerrung aufgrund von Eigeninteressen, Erinnerungslücken oder persönlichen Voreingenommenheiten. Bei der Interpretation von Bewertungen aus Sicht der Forschung besteht die Gefahr, dass der Forscher ungewollt seine Bewertungsordnung zugrunde legt und dadurch den intendierten Handlungssinn verfehlt. Dementsprechend ist in beiden Fällen darauf zu achten, dass Auskünfte zu Auswahl und Gewichtung von Kriterien nicht als unumstößliche „objektive Wahrheiten“ fehlgedeutet werden.

Viertens kann die Analyse aufgrund der Beschränkung auf einen Messzeitpunkt (das Jahr 2015) nur eine Momentaufnahme bieten. Für eine Klärung der Frage, wie sich die Kriterien an Fachhochschulen und Universitäten im Bereich der professoralen Personalauswahl *entwickeln*, ist eine Längsschnittuntersuchung unabdingbar. Die vorliegende Studie könnte als Referenzpunkt für eine solche Untersuchung wie auch als Anregung für Forschungsaktivitäten zu anderen Aspekten von Berufungsverfahren an staatlichen Hochschulen in Deutschland dienen.

Literatur

Abele-Brehm, A.; Bühner, M. (2016): Wer soll die Professur bekommen? Eine Untersuchung zur Bewertung von Auswahlkriterien in Berufungsverfahren der Psychologie. Psychologische Rundschau, 67(4), 250–261

Auspurg, Katrin; Hinz, Thomas; Schneck, Andreas (2017): Berufungsverfahren als Turniere: Berufungschancen von Wissenschaftlerinnen und Wissenschaftlern. Zeitschrift für Soziologie, 46(4), 283–302

Baier, Christian; Münch, Richard (2013): Institutioneller Wettbewerb und Karrierechancen von Nachwuchswissenschaftlern in der Chemie. KzFSS Kölner Zeitschrift für Soziologie und Sozialpsychologie, 65(1), 129–155

Bauer, Martin; Wolff, Cathrin (2017): TeilnehmerInnen im Fokus. Ergebnisse einer Analyse der Teilnehmendenstruktur an der Schnittstelle zwischen universitärem System und dem Weiterbildungsmarkt. In: Egger, Rudolf; Bauer, Martin Heinz (Hrsg.): Bildungspartnerin Universität. Tertiäre Weiterbildung für eine erfolgreiche Zukunft. Wiesbaden: Springer VS, S. 75–94

Blasius, Jörg (1994): Empirische Lebensstilforschung. In Lebensstile in den Städten (pp. 237–254). Wiesbaden: VS Verlag für Sozialwissenschaften, S. 237–254

Blasius, Jörg (2001): Korrespondenzanalyse. München/Wien: Oldenburg

Blasius, Jörg (2010): Korrespondenzanalyse. In: Wolf, Christof; Best, Henning (Hrsg.): Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: Springer VS, S. 367–389

Bloch, Roland/Lathan, Monique;Mitterle, Alexander;Trümpfer, Doreen;Würmann, Carsten (2014): Wer lehrt warum? Strukturen und Akteure der akademischen Lehre an deutschen Hochschulen. Leipzig: Akademische Verlagsgesellschaft

Borgwardt, Angela (2016): Zwischen Forschung und Praxis. Die Rolle der Fachhochschulen im Wissenschaftssystem. Berlin: Friedrich-Ebert-Stiftung

De Boer, Harry F; Enders, Jürgen;Schimank, Uwe (2008): Comparing higher education governance systems in four European countries. In: Soguel, Nils C. u.a. (Hrsg.): Governance and Performance of Education Systems. Dordrecht, S. 35–54

DiMaggio, Paul J.;Powell, Walter W. (1983): The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. In: American Sociological Review. 48(2), 147–160

Duong, Sindy; Hachmeister, Cort-Denis;Roessler, Isabel (2014): Gleichzeitigkeit des Ungleichzeitigen? Lage und Zukunft von Fachhochschulen im Hochschulsystem aus Sicht von Fachhochschulleitungen. CHE-Arbeitspapier Nr. 180. Gütersloh: Centrum für Hochschulentwicklung

Hachmeister, Cort-Denis; Duong, Sindy;Roessler, Isabel (2015): Forschung an Fachhochschulen aus der Innen- und Außenperspektive: Rolle der Forschung, Art und Umfang. CHE-Arbeitspapier Nr. 181. Gütersloh: CHE

Hartmer, Michael; Detmer, Hubert (2017): Hochschulrecht. Ein Handbuch für die Praxis. 3., neu bearbeitete Auflage. Heidelberg: C.F. Müller

HLB (Hochschullehrerbund) (o.J. [2013]): Online-Befragung zur Workload von Professorinnen und Professoren an Fachhochschulen. Durchgeführt vom Mittelstands-Institut an der Hochschule Kempten e.V. (https://hlab.de/fileadmin/_migrated/news_uploads/2013-hlb-Workload-Umfrage-Ergebnisse-FINAL.pdf. Zuletzt abgerufen am 27.9.2017)

Huber, Michael (2012): Die Organisation Universität. In: Maja Apelt;Veronika Tacke (Hrsg.): Handbuch Organisationstypen. Wiesbaden: Springer VS, 239–252

In der Smitten, Susanne; Sembritzki, Thorben;Thiele, Lisa;Kuhns, Johannes;Sanou, Amadou;Valero-Sanchez, Marco (2017): Bewerberlage bei Fachhochschulprofessuren (BeFHPro). Forum Hochschule Heft 3. Hannover: Deutsches Zentrum für Hochschul und Wissenschaftsforschung. (www.dzhw.eu/pdf/pub_fh/fh-201703.pdf. Zuletzt abgerufen am 27.9.2017)

In der Smitten, Susanne; Sembritzki, Torben; Thiele, Lisa (2017a): Bewerberlage bei Fachhochschulprofessuren – Unzureichend strukturierte Karrierewege erschweren die Stellenbesetzung. (DZHW Brief 1|2017). Hannover: DZHW

In der Smitten, Susanne; Sembritzki, Torben; Thiele, Lisa (2017b). Schwierigkeiten bei der Besetzung von Fachhochschulprofessuren. Die Neue Hochschule, Heft 5(2017), 26–30

Klawitter, Maren (2017): Die Besetzung von Professuren an deutschen Universitäten. Empirische Analysen zum Wandel von Stellenprofilen und zur Bewerber(innen)auswahl. Dissertation Universität Kassel (Im Erscheinen)

Kleimann, Bernd; Klawitter, Maren (2017): An Analytical Framework for Evaluation-Based Decision-Making Procedures in Universities. In: Jeroen Huisman ; Malcolm Tight (ed.): Theory and Method in Higher Education Research. Volume 3. Bingley: Emerald Publishing, S. 39–57

Krücken, Georg; Meier, Frank (2006): Turning the University into an Organizational Actor. In: Gili S. Drori; John W. Meyer; Hoky Hwang (Hrsg.): Globalization and Organization. World Society and Organizational Change. Oxford: Oxford University Press, 240–257

Kühl, Stefan (2011): Organisationen. Eine sehr kurze Einführung. Wiesbaden: Springer

Le Roux, Brigitte; Rouanet, Henry. (2004): Geometric Data Analysis. Dordrecht: Springer

Le Roux, Brigitte; Rouanet, Henry. (2010): Multiple Correspondence Analysis. Sage

Luhmann, Niklas (2006): Organisation und Entscheidung. 2. Auflage. Wiesbaden: Springer

Meier, Frank (2009): Die Universität als Akteur. Zum institutionellen Wandel der Hochschulorganisation. Wiesbaden: VS Verlag für Sozialwissenschaften

Möller, Christina (2015): Herkunft zählt (fast) immer. Soziale Ungleichheiten unter Universitätsprofessorinnen und -professoren. Weinheim: Beltz

Mundt, Fabian; Kutzner, Daniela (2017): Die Konstruktion des Raumes der Interessenlagen potenzieller Studierender in der Frühpädagogik. Die hochschullehre, Jahrgang 3/2017

Preißler, Ulrike (2016): Die W-Besoldung in 17 Variationen. In: Forschung & Lehre, Heft 4, S. 316–319

Rössel, Jürgen, Roose, Jochen (Eds.). (2015): Empirische Kulturosoziologie: Festschrift für Jürgen Gerhards zum 60. Geburtstag. Springer VS

Schiller, Nadine; Mahmud, Fereshta; Kenkel, Eva (2015): Factsheet Fachhochschulen und Universitäten. Ein Vergleich auf Basis von statistischen Kennzahlen. [Gütersloh:]

Schimank, Uwe (2014): Der Wandel der „Regelungsstrukturen“ des Hochschulsystems und die Folgen für die wissenschaftliche Forschung – zur Entwicklung der Governance-Perspektive. In: Bora, Alfons u. a. (Hrsg.): Wissensregulierung und Regulierungswissen. Weilerswist, S. 19–40

Schmid, Christian J.; Wilkesmann, Uwe. (2015): Ansichtssache Managerialismus an deutschen Hochschulen – ein empirisches Stimmungsbild und Erklärungen. Beiträge zur Hochschulforschung, 37(2), 56–87

Schnell, Lisa (2013): Streit um den Doktorhut. In: taz. Die tageszeitung vom 4.12.2013, S. 18

Schreiterer, Ulrich (2016): Die Folle der Fachhochschulen im Wissenschaftssystem. In: Borgwardt (2016), S. 19–29

Simon, Herbert (1981 [1945]): Entscheidungsverhalten in Organisationen. Landsberg am Lech: Verlag Moderne Industrie

Statistisches Bundesamt (2016a): Bildung und Kultur. Monetäre hochschulstatistische Kennzahlen 2014. Fachserie 11. Reihe 4.3.2. Wiesbaden

Statistisches Bundesamt (2016b): Personal an Hochschulen – vorläufige Ergebnisse. Reihe „Bildung und Kultur“. 1. Dezember 2016. (https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Hochschulen/PersonalVorbericht5213402168004.pdf?__blob=publicationFile) Zuletzt abgerufen am 12.02.2018

Vogel, Michael P. (2009): The professionalism of professors at German Fachhochschulen. In: Studies in Higher Education, 34(8), 1–16

Walgenbach, Peter; Meyer, Renate E. (2008): Neoinstitutionalistische Organisations-
theorie. Stuttgart: W. Kohlhammer

Wissel, Carsten von (2007): Hochschule als Organisationsproblem. Neue Modi universitärer Selbstbeschreibung in Deutschland. Bielefeld: Transcript Verlag

Wissenschaftsrat (2010): Empfehlungen zur Rolle der Fachhochschulen im Hochschulsystem. Drs. 10031–10. Berlin. (<https://www.wissenschaftsrat.de/download/archiv/10031-10.pdf>. Zuletzt abgerufen am 27.9.2017)

Würmseer, Grit (2010): Auf dem Weg zu neuen Hochschultypen. Eine organisationssoziologische Analyse vor dem Hintergrund hochschulpolitischer Reformen. Wiesbaden: VS Verlag für Sozialwissenschaften

Anhang

Tabelle 1: Beiträge der Kategorien zur Varianz der 1. Dimension

-	durchgeführte Forschungsprojekte	9,336310862
-	Forschungskonzept	7,607426913
+	Betreuung des wiss. Nachwuchses	6,407286823
+	internationale Ausrichtung der Bewerberin	5,998339077
+	Mitwirkung in der Scientific Community	5,610290240
+	eingeworbene Drittmittel	5,589232526
-	wissenschaftliche Publikationen	5,395359355
+	Kooperationen mit Hochschulen	5,350126544
-	internationale Ausrichtung der Bewerberin	4,209360756
+	durchgeführte Forschungsprojekte	4,180437700
-	Kooperationen mit Hochschulen	3,835046461
+	Forschungskonzept	3,573185368
-	eingeworbene Drittmittel	3,446693391
-	Betreuung des wiss. Nachwuchses	3,303757268
-	wissenschaftlicher Werdegang	2,880766226
+	Mitwirkung an Studiengängen und Lehrangeboten	2,670269151
+	Kooperationen mit Profit- und Non-Profit	2,566532645
-	Mitwirkung in der Scientific Community	2,512070257

Tabelle 2: Beiträge der Kategorien zur Varianz der 2. Dimension

-	Lehrkonzept	11,89295601
-	Lehrvortrag im Berufungsverfahren	10,73458466
-	Lehrerfahrung	7,770816658
+	Lehrevaluationsergebnisse	6,093669715
-	Lehrevaluationsergebnisse	5,496251115
-	besondere fachspezifische Anforderungen	4,658253726
-	Passung der Persönlichkeit	4,376987588
-	Fachvortrag im Berufungsverfahren	4,142651915
+	besondere fachspezifische Anforderungen	3,632583181
-	Passung mit Rahmenbedingung	3,279990751
+	Lehrkonzept	3,187001284
-	wissenschaftliche Publikationen	3,073194418
+	Mitwirkung in Gremien und Kommissionen	2,769772447
+	Teilnahme an hochschuldidaktischer Qualifizierung	2,671330877

Tabelle 3: Erklärte Varianz pro Achse

	%	kumuliert in %
1. Dimension	16,1	16,1
2. Dimension	13,9	30,0
3. Dimension	7,0	37,0
4. Dimension	5,6	42,6
5. Dimension	4,8	47,4
6. Dimension	4,7	52,1

Tabelle 4: η^2 der Dimensionen 1–5

	1. Dimension	2. Dimension	3. Dimension	4. Dimension	5. Dimension
Organisation	20,1	19,5	27,0	0,0	4,6
Fach	9,5	9,4	6,7	3,1	0,6

Artikel eingereicht: 29.09.2017
Artikel angenommen: 28.02.2018

Anschrift der Autoren:

PD Dr. Bernd Kleimann

Malte Hückstädt

Deutsches Zentrum für Hochschul- und Wissenschaftsforschung GmbH (DZHW)

Lange Laube 12

30159 Hannover

E-Mail: kleimann@dzhw.eu

hueckstaedt@dzhw.eu

Bernd Kleimann ist kommissarischer Leiter der Abteilung „Governance in Hochschule und Wissenschaft“ des Deutschen Zentrums für Hochschul- und Wissenschaftsforschung (DZHW) in Hannover.

Malte Hückstädt ist Masterstudent der Soziologie mit dem Schwerpunkt quantitative Methoden an der Universität Bielefeld und studentische Hilfskraft am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) in Hannover.

Measuring “Teaching Excellence” and “Learning Gain” in the United Kingdom

Heike Behle, Sunil Maher

In the United Kingdom’s (UK) higher education system, teaching and learning are currently under review to evaluate both social and individual investment to tertiary education. The UK government introduced the Teaching Excellence Framework and Student Outcomes Framework to assess the quality of undergraduate teaching in universities and other higher education providers. There are 14 research projects which are identifying multiple ways to measure Learning Gain, a key component of the Teaching Excellence Framework and Student Outcomes Framework. Teaching Excellence encompasses Teaching Quality, the Learning Environment as well as Student Outcomes and Learning Gain, it uses both core metrics and a narrative to evaluate individual universities performance. Current research evaluate methodologies to measure Learning Gain and determine the potential suitability and scalability of possible measures across the sector. This paper describes the current state of policy and research against the specific background of the UK higher education sector.

1 Introduction: Measuring Learning Gain in Higher Education

Measurement of higher education outcomes is a part of the political agenda in many countries and international organisations. For example, the OECD carried out a feasibility study for the Assessment of Higher Education Learning Outcomes (AHELO), aimed to test methods to assess students’ level of knowledge and skills graduation, and in the USA, the Wabash National Study focused on teaching and learning in liberal arts education. However, while in some countries there is a huge effort to identify indicators to measure learning gain across different institutional settings and in different subjects this is not the case in most other countries. Currently, the level of research on learning gain in various other countries has been described as not sufficient, and most previous studies concentrate on local settings with a small group of participants to determine prerequisites for admissions tests, gather data on learning opportunities, and use self-report measures of outcomes (*Zlatkin-Troitschanskaia, Shavelson & Kuhn, 2015*).

The engagement with higher education outcomes is usually justified as an exercise to understand if the tertiary education sector is providing the right kind of skills and knowledge to cater for the future. It is the intention “to recognise excellent undergraduate teaching and learning in universities and colleges” (*Beech, 2017*). Another

reason is the lack of comparability of classes of degrees across different higher education institutions and across different subjects and courses (Arum & Roksa, 2011; Zlatkin-Troitschanskaia, Shavelson and Kuhn, 2015). One particular aspect in this regard is the identification of skills and knowledge needed in a future labour market.

Policy-makers in the UK are concerned with allocation of existing resources to "produce" the skills, knowledge and attitudes future employees require (Clarke, 2017). In order to do so, it will be required to understand what it is that a tertiary study does, how individuals and the society benefits from it, and crucially, how excellence in teaching and learning can be identified. This knowledge is required not only for policy makers and employers but also for potential applicants to higher education. In the shift from elite education to a highly diverse provision of a mass higher education system, graduates need to be able to identify and document the skills, knowledge and competences they acquired during their tertiary studies. Methodologically, this results in a shift in the assessment of students' achievement. *"Higher education globally has moved from an era of elite intake and stability of cohorts in which assessment was used to judge students against each other, to one that is slowly becoming outcomes-based, using explicit standards against which each student is judged"* (Boud, 2017).

In the UK, the focus is mainly on *learning gain* which compares performance at two points in time, as opposed to *learning outcome* (Schleicher 2015) which concentrates exclusively on the output level. *Learning gain* in higher education measures the "distance travelled" with regards to skills, competencies, content knowledge and personal development demonstrated by students at two points in time (Hoareau McGrath et al., 2015). The "distance-travelled model" needs to be differentiated to the "value-added model". The first compares between two measures of actual student performance, while the latter is based on the comparison between performance predicted at the outset of studies and actual performance achieved (Hoareau McGrath et al., 2015). Initially, these points in time refer to the start of higher education and graduation, however, a long-term approach could also account for further development after graduation.

This paper reports on current policy in England¹ to measure teaching excellence, as an example of ongoing attempts to measure learning gain. We describe the procedure to measure teaching excellence, the so-called Teaching Excellence and Student Outcomes Framework, and on-going research to measure learning gain. The UK has the most marketised higher education system within Europe, with students as consumers who see higher education as a direct investment which renders the outcome of higher education to potential achievements on the labour market (Tomlinson, 2017). This

¹The Teaching Excellence and Student Outcomes Framework is for England only. Wales, Scotland and Northern Ireland have devolved governments but can participate on a voluntary basis.

human capital approach has exacerbated the urgent need for (potential) students, as investors, to understand their expected return. Furthermore, it has sharpened the need to understand the provision of teaching and learning which has resulted in the implementation Higher Education Research Act (2017) and the Teaching Excellence and Student Outcomes Framework.

This paper will first introduce the specialities of the UK Higher Education Sector, followed by a description of the Teaching Excellence and Student Outcomes Framework. A third section will introduce both the themes covered in the learning gain projects and the methodological considerations followed by an early conclusion.

2 The Higher Education Sector in the United Kingdom

Higher education in the United Kingdom (UK) can be described as a marketised higher education system, characterised by three features:

- (1) Relatively high tuition fees in a European context: Currently, UK domiciled students entering in higher education in 2017 are charged up to £9,250 annually, depending on their domicile status and the higher education provider², with international students charged considerably more. The level of tuition fees are politically contested and it is unclear whether tuition fees will remain at current levels.
- (2) An exceedingly diversified network of providers: Applicants to higher education can choose their course providers from a much diversified higher education system ranging from the highly reputable Russell Group Universities to smaller Further Education Colleges with degree awarding powers. A wide variety of different higher education providers are expected to cater for a highly diverse student population and the current government plans to further diversify it.
- (3) Customer-orientated students: Finally, tuition fees have changed students' attitude to their studies, their choice of courses and the providers. A previous study found that students have become more instrumental in their choice of courses and many were likely to give an employment-related main reason for applying to higher education at the start of their courses (*Purcell et al 2008:35ff*). Higher education courses are marketed as investment commodities (*Naidoo and Jamieson 2005*), and higher education is increasingly seen as an investment into future employment opportunities.

²At today's rate, £9,250 are equal to approximately €10,500. A breakdown of the different levels of tuition fees according to domicile and location of study can be found here: <https://www.ucas.com/ucas/undergraduate/finance-and-support/undergraduate-tuition-fees-and-student-loans> (accessed 2017-10-01)

Despite the high labour-market orientation of students and a high rate of employment after graduation *per se*, many graduates work in non-graduate jobs for which a degree is not necessary (Elias & Purcell, 2004). A recent study (Behle, 2016) found that of all employed graduates, approximately 30 per cent from three year courses and approximately 40 per cent from four year courses worked in non-graduate employment indicating that higher education remains a high risk financial investment and higher education studies cannot guarantee a graduate job (Brynin 2013). Employment in non-graduate jobs, however, is more likely for graduates from specific subjects, lower access providers and regional location (Behle, 2016). The graduate premium which measures the difference between graduates' achieved salary and that of someone who could have attended higher education but choose not to, exists but the growth of earnings of graduates in their early careers has slowed down in recent years relative to the increase in average earnings across the UK economy (Purcell *et al.*, 2013).

In this contradictory context, students' high expectations for a career in their chosen career path after graduating with a large debt and, at the same time, a high proportion of graduates working in non-graduate jobs has resulted in a revision of teaching and learning in higher education, and in 2016, the government implemented the Teaching Excellence and Student Outcomes Framework to measure the differences in performance between higher education providers. At approximately the same time, following findings from the USA (Arum & Roksa, 2011) where authors found a lack of progress in the critical thinking of students, the Higher Education Funding Council for England commissioned research to pilot and evaluate a range of approaches for measuring learning gain. This paper describes the Teaching Excellence Framework and Student Outcomes Framework and the work of various research projects to measure Learning Gain in Higher Education in the UK. It is important to note that the Teaching Excellence and Student Outcomes Framework, being new legislation, is undergoing further development, and this paper, written in autumn 2017, reviews the previous assessment of Teaching Excellence and Student Outcomes Framework 2 and governmental guidance referring to the upcoming Teaching Excellence Framework and Student Outcomes Framework 3.

3 The Teaching Excellence and Student Outcomes Framework (TEF)

The Teaching Excellence and Student Outcomes Framework (TEF) was introduced in the academic year 2015/2016 to provide students with additional information about the kind or level of teaching and learning they can expect from a provider of tertiary education. It was branded as a "*new scheme for recognising excellent teaching, in addition to existing national quality requirements for universities, colleges and other higher education providers*" (Higher Education Funding Council for England website). At the current stage (October 2017), the TEF is voluntary, covers undergraduate teach-

ing, and measures performance at an institutional level only, however, it is anticipated that it will become compulsory and will also cover the subject level. The TEF aims to help students to be better informed and find information about what and where to study, to raise esteem for teaching, to recognise and reward excellent teaching, and to narrow the gap between the skills, knowledge and competencies students gain and the needs of employers, business, industry and the professions (*Department for Education, 2017*). The TEF includes all undergraduate provision of higher education as well as degree apprenticeships or integrated Master's degrees irrespective of delivery mode (such as part-time, or as distance learning course).

The TEF assessment framework has been developed to take into account diverse forms of teaching and learning excellence. UK universities are traditionally divided into pre-1992 universities, which tend to be more research-oriented, and post-1992 universities, often former polytechnic colleges, which gained their university status in a national effort to widen participation in degree-level education. However, further divisions exist, for example between the 24 research-intensive older universities (the Russell Group) and other groups (such as the Million+ or the University Alliance), with some universities concentrating solely on specific courses only, for example in Arts or Medicine. Additionally, it should be noted that in the application process, universities select students according to their achievements at secondary schools which further diversify institutions according to their student population.

3.1 Assessment criteria

The Teaching Excellence Framework and Student Outcomes Framework (TEF) concentrates on three aspects of quality. "Student Outcomes and Learning Gain", "Teaching Quality" and the "Learning Environment". "Student Outcomes and Learning Gain" measures the achievement of positive outcomes, including "*acquisition of attributes such as lifelong learning skills and others that allow a graduate to make a strong contribution to society, economy and the environment*". It also measures "*progression to further study, acquisition of knowledge, skills and attributes necessary to compete for a graduate level job that requires the high level of skills arising from higher education*" (*Department for Education, 2017, p. 24*). "Teaching Quality" includes the provision of different forms of structured learning to allow students to contact, stimulate and challenge and to encourage student engagement and effort. The "Learning Environment" refers to existence of resources such as libraries, laboratories and design studios, work experience, opportunities for peer-to-peer interaction and extra-curricular activities in supporting students' learning and the development of independent study and research skills. The different criteria to assess the three aspects of quality are outlined in Table 1, below.

Table 1: Teaching Excellence Framework and Student Outcomes Framework assessment criteria

Aspect of Quality Areas of teaching and learning quality	Reference	Criterion
Teaching Quality	Student Engagement	Teaching provides effective stimulation, challenge and contact time that encourages students to engage and actively commit to their studies
	Valuing Teaching	Institutional culture facilitates, recognises and rewards excellent teaching
	Rigour and Stretch	Course design, development, standards and assessment are effective in stretching students to develop independence, knowledge, understanding and skills that reflect their full potential
	Feedback	Assessment and feedback are used effectively in supporting students' development, progression and attainment
Learning Environment	Resources	Physical and digital resources are used effectively to aid students' learning and the development of independent study and research skills
	Scholarship, Research and Professional Practice	The learning environment is enriched by student exposure to and involvement in provision at the forefront of scholarship, research and/or professional practice
	Personalised Learning	Students' academic experiences are tailored to the individual, maximising rates of retention, attainment and progression
Student Outcomes and Learning Gain	Employment and Further Study	Students achieve their educational and professional goals, in particular progression to further study or highly skilled employment
	Employability and Transferable Skills	Students acquire knowledge, skills and attributes that are valued by employers and that enhance their personal and/or professional lives
	Positive Outcomes for All	Positive outcomes are achieved by its students from all backgrounds, in particular those from disadvantaged backgrounds; or those who are at greater risk of not achieving positive outcomes

Source: Adapted from Department for Education (2017), Teaching Excellence and Student Outcomes Framework Specification, p. 25/26.

A key element of all aspects of quality is the extent to which positive outcomes are achieved for all students, including those from disadvantaged backgrounds such as mature or disabled students, those from Black or Minority Ethnic groups, or other aspects leading to disadvantages such as family background. In England, universities are legally required to invest in activities designed to widen the participation for disadvantaged and under-represented groups, often abbreviated as WP (Widening Participation) (see for example *Harrison & Waller, 2017*).

All Teaching Excellence Framework and Student Outcomes Framework (TEF) submissions from the second assessment (TEF 2) are now available online³, with regards to student outcomes and learning gain. The following examples illustrate how narrative submission were evidenced, these reflect different higher education institutions.

"[University] takes pride in being an inclusive and diverse community. (...) This diversity allows all students to encounter, be challenged by, and learn from a wide range of views and experiences."

"Students often enter [University] with comparatively low social capital."

Once the level of diversity is established, the outcome for graduates and specific ways to enhance students' employability are described, as these examples show.

"Over 70 % of [University] students take a placement in the third year of their course, or engage in the equivalent through clinical practice. (...) Every student on placement has a contract or learning agreement with their employer, and job descriptions are jointly agreed with employer and the University. All students are required to keep a reflective workbook, as well as submitting an academic assignment based on their experiences."

"[University] students develop significant transferable skills during their studies. In supervisions students develop the ability to engage in an academic discussion and debate with an expert on a regular basis, producing skills that employers value highly, such as the ability to formulate and respond to arguments, assimilate complex ideas and information rapidly and accurately, and apply information in new contexts."

"Many [University] students are very actively involved in enterprise and entrepreneurship through clubs and societies."

"Support and advice to students is provided by our Careers Service, which is one of the largest and best resourced in the UK, providing students and graduates up to two years post-graduation with access to thousands of career opportunities, both in the UK and overseas."

"Employability modules are embedded in many of our undergraduate programmes, and over a fifth of our undergraduates undertake one or more of these modules while studying with us (...) These modules address 'readiness to work' and 'employability skills' in a focused way, supporting a higher proportion of successful outcomes for the students engaging with them."

The narrative supplements the achieved core metrics score of teaching excellence, as described in the next chapter.

³<http://www.hefce.ac.uk/lt/tef/data/>

3.2 Core metrics

The Core Metrics to assess the three aspects of quality use data from the National Students’ Survey and on students and graduates collected by the Higher Education Statistics Agency.

Table 2: Core Metrics Assessment Framework

Aspect of Quality Areas of teaching and learning quality	Teaching Quality	Learning Environment	Student Outcomes and Learning Gain
Core Metrics	Teaching on my course (NSS) Assessment and feed- back (NSS)	Academic support (NSS) Continuation (HESA)	Employment/ further study (DLHE) Highly-skilled employment/ further study (DLHE)

Source: Adapted from Department for Education (2017), Teaching Excellence and Student Outcomes Framework Specification, p. 23. The surveys on which the core metrics are based on are the National Students’ Survey (NSS), Higher Education Statistics Authority (HESA), and the Destination of Leavers of Higher Education (DLHE) survey.

Introduced in 2005, the National Students’ Survey (NSS) was aimed at final year students, it seeks students’ views on various aspects of teaching, assessment and support provided by their university its courses, and ultimately to produce league tables of university performance (*Douglas et al., 2006*). According to the NSS website⁴, nearly three million students have responded to the NSS which represents more than 70 per cent of all final-year students in 2016. The NSS asks students to respond to 27 statements using Likert scales from “definitely agree” to “definitely disagree”, with a “not applicable” option.

For the Teaching Excellence and Student Outcomes Framework, the following statements were identified as relevant to teaching quality and the learning environment (2016 NSS survey⁵):

■ Statements to “The teaching on my course”

- Staff are good at explaining things.
- Staff have made the subject interesting.
- The course is intellectually stimulating.

⁴More information about the NSS can be found here www.thestudentsurvey.com

⁵The NSS questions used from 2005–2016 were published here www.bristol.ac.uk/academic-quality/ug/nss/nssqs05-16.html/ (accessed 2017-11-28)

■ **Statements to “Assessment and feedback”**

The criteria used in marking have been made clear in advance.

Assessment arrangements and marking have been fair.

Feedback on my work has been prompt.

I have received detailed comments on my work.

Feedback on my work has helped me clarify things I did not understand.

■ **Statements to “Academic support”**

I have received sufficient advice and support with my studies.

I have been able to contact staff when I needed to.

Good advice was available when I needed to make study choices.

The Higher Education Statistics Authority holds both student records and organises the Destination of Leavers of Higher Education (DLHE)⁶ survey.

The Higher Education Statistics Agency collects personal data together with information about their courses from all UK Higher Education Institutions forming an individual “Student Record”, the Teaching Excellence Framework and Student Outcomes Framework (TEF) includes continuation/non-continuation rate as part of the framework. The current DLHE is a survey of graduates and focusses on students’ activities six months after leaving university. Both general employment and further study are included in the Teaching Excellence Framework and Student Outcomes Framework calculations.

Supplementary metrics, introduced for Teaching Excellence and Student Outcomes Framework 3, refer to the Longitudinal Education Outcomes dataset, to describe the sustained employment or further study or the above median earnings threshold. The LEO dataset links students’ personal data and information about their education with employment and income, and any benefits claimed using administrative data from the tax office and the Department for Work and Pensions. In addition, contextual data on the body of students such as level of study, age, gender, ethnicity, disability, entry qualification or the domicile can also be taken into account for the TEF assessment.

3.3 Benchmarking and assessing

In contrast to existing quality assurance and ranking exercises, and in the context about the specific mix of students and subjects taught at that provider, the Teaching Excellence Framework and Student Outcomes Framework (TEF) uses benchmarks to allow interpretation of a provider’s metrics. For each provider and each core metric, an

⁶The current DLHE will be replaced in the next years with a new survey named ‘Graduate Outcomes’, where graduates will be surveyed approximately 15 months after completing their studies.

individual benchmark is calculated. *"The benchmark is a weighted sector average where weightings are based on the characteristics of the students at the provider"* (Department for Education, 2017, p. 38/39). The benchmarking prevents a comparison between unequal types of universities, for example, the performance of a vocationally-orientated former polytechnic college with a large number of students with lower achievements at secondary schools cannot be compared with a high reputable university which attracts and selects only the best students.

Each provider will be "flagged" in the specific core metric, where the results are at least two percentage points above or below its benchmark and that difference is statistically significant. The positive or negative differences are marked by "+" and "-", causing different results ("+" for 2 percentages above the benchmark, "++" for 3 or more percentages above the benchmark, with "-" for 2 percentages below the benchmark, "--" for 3 or more percentages below the benchmark). Based on this flagging exercise, for each provider the amount of positive and negative flags are calculated as follows; "Teaching on my course", "Assessment and feedback" and "Academic support" are entered with a weight of 0.5, and the other core metrics are weighted by a 1.0. The method of calculation was not used for the TEF 2, but will be implemented for TEF 3.

"A provider with positive flags (either "+" or "++") in core metrics that have a total value of 2.5 (...) or more and no negative flags (either "-" or "--") should be considered initially as Gold. A provider with negative flags in core metrics that have a total value of 1.5 or more (...) should be considered initially as Bronze, regardless of the number of positive flags. All other providers, including those with no flags at all, should be considered initially as Silver." (Department for Education, 2017, p. 58). This initial rating is then adjusted depending on the University's performance in "split metrics" where the same measures were split among different groups of students according to characteristics like gender, ethnicity and social background. Finally, the results are reviewed against Universities' narrative submissions and contextual evidence provided.

For example, a provider submitted figures based on the three metrics based on the aspects of quality. These metrics, for each aspect of quality, get individually benchmarked with those of a similar provider (e.g. according to the proportion of students from a poorer background). Flags are then allocated in all core metrics according to the difference between the specific provider and the benchmark. Finally, the amount of flags across the core metrics are summed up, and the achieved total will result in a Gold, Silver or Bronze award.

3.4 Reception of the Teaching Excellence Framework and Student Outcomes Framework 2

The Teaching Excellence Framework and Student Outcomes Framework (TEF) is currently being coordinated by Higher Education Funding Council for England but will transfer to the Office for Students (OfS) on 1st April 2018. From October 2017, and alongside provider delivery TEF, subject-level TEF is now being piloted.

A “Gold” rating is evidence of provision of outstanding and highest quality in the UK Higher Education sector and was received by 45 colleges and universities. “Silver” indicates high quality and consistently exceeding the national quality requirements for UK, with 67 Higher Education Institutions were awarded with this rating. 25 institutions were awarded “Bronze” which means that the provision meets the national quality requirements for UK Higher Education (*Department for Education, 2017; Times Higher Education, 2017*). While in many cases the TEF repeated previous rankings and many highly reputable institutions received a “Gold” standard, some world-renowned universities received a bronze award indicating that they might have prioritised excellence in research over teaching and student experience⁷. An appeal is only permitted on a very narrow set of grounds (i.e. formality) and excluded for the nature of the rules and processes. A few institutions legally challenged their TEF award, however only one University (University of East Anglia) successfully upgraded their “Silver” to a “Gold” award.

One of the critique points refers to the lack of knowledge about implications of the TEF, and many expect that there will be a connection between the TEF results and institutions’ opportunities to raise tuition fees. While the linkage between TEF and tuition fees were removed by an amendment from the House of Lords⁸, students’ boycott of the National Students’ Survey, organised by Students’ union bodies, went ahead⁹. As described above, the results are partially based on direct responses from students (“Teaching on my course”, “Assessment and feedback” and “Academic support”), for TEF 3 the weighting of the student-led responses has been reduced by 50 per cent, this has been widely criticised as diminishing the student voice¹⁰.

4 Learning Gain in Higher Education – ongoing research

In an evaluation of the Teaching Excellence Framework and Student Outcomes Framework (TEF) (Beech, 2017), Chris Husbands, the first TEF panel Chair, acknowledges

⁷<http://wonkhe.com/blogs/tef-results-how-do-ref-and-tef-results-compare/>

⁸<http://wonkhe.com/blogs/government-defeated-in-the-lords-over-tef-and-fees/>

⁹<https://www.nusconnect.org.uk/articles/the-time-to-act-is-now-boycott-the-nss/>

¹⁰<http://wonkhe.com/blogs/halving-the-nss/>

that "*the TEF will not stand still*" (p. 3). It is anticipated that current research to measure learning gain in higher education will be used to revise the TEF and to fine-tune existing instruments. Across the Higher Education sector, many research projects are currently engaged in developing new measures of learning gain in order to better evidence student outcomes of the TEF. Learning gain, in this context, is defined as the "distance travelled" from higher education entry through to graduation, or the improvement in knowledge, skills, work-readiness and personal development demonstrated by students at two points in time (Hoareau McGrath et al., 2011). A specific interest, in this respect, is the development of students' employability, defined as the ability to find, retain and progress in employment (Green et al., 2013).

The Learning Gain Pilot Programme¹¹, funded by Higher Education Funding Council for England, invited English Higher Education Institutions and Further Education Colleges to submit proposals to pilot and evaluate methodologies to measure learning gain and determine potential validity, suitability and scalability¹² of possible measures across the sector. A total of 13 pilot projects¹³ were awarded funding to investigate existing or new methods. The projects use different methodologies in their approaches to investigate this area of work, more specifically a range of grades, surveys, standardised tests, other qualitative methods or mixed methods approaches are used to further their investigations. Some proposed methods account for psycho-social, behavioural, and cognitive components and explore study abroad, early intervention, student engagement, self-efficacy, critical thinking and virtual learning environments.

The following section will discuss two aspects from the Learning Gain Pilot Programme, different themes to students' learning and employability gain, and methodological considerations.

4.1 Themes covered in the Learning Gain projects

Most pilot projects have identified various themes which will need to be considered for wide-scale implementation. Learning gain projects are conducted by research-orientated as well as vocationally-orientated Higher Education Institution's, and additionally, some specialist Higher Education Institution's such as Conservatoires of Music and Dance or the distance-learning Open University are included as the Higher Educa-

¹¹<http://www.hefce.ac.uk/lt/lg/>

¹²Scalability, in this context, is seen as wide application across England's HE Sector. Moreover, it explores the "appropriateness, practicality and value for money of extending the use of measures of learning gain" (Kandiko Howson, 2018, p. 9). The scalability of metrics used to measure the outcome is a particular challenge as the HE sector is diverse. Therefore, scalability refers to the comparability of Learning Gain across disciplines and HE Institutes.

¹³<http://www.hefce.ac.uk/lt/lg/projects/>

tion Founding Council for England intended to represent the diversity of students and the Higher Education sector.

The engagement with students at an early stage of entry into higher education could provide a baseline for students' level of skills, knowledge and attitudes, it could also be used as a proxy for their preparedness to study at a higher level. For example, the Careers Registration project led by the University of London has partnered with 16 Higher Education Institutions, it asks students to respond to set questions, integrated as part of the annual registration process, to determine their career and work readiness. Students are requested to select pre-determined statements to best describe the attitudes and actions taken on their career aspirations and possible employment experience. Such an approach allows other functional areas of institutes, such as the career service to gauge students' expectations and inform planning, and other departments, of their needs, expectations or requirements. The net effect will be development, creation or enhancement of services that will (further) support and inform students of possible opportunities and, importantly, how to realise them. The provision of such data will identify socio-economic, "at risk" and Widening Participation groups that require more prescribed support services.

The assessment of both transferable and subject-related skills and knowledge together with students' strengths and weaknesses by self-assessment or psychometric tests is the focus of other projects who will further track the development over the time. Whilst psychometric tests are predominantly used within an employer's recruitment process to determine an individual being a "good fit" to a working environment, there are advantages to students possessing such self-awareness at an earlier stage, potentially assisting them in identifying any development needs. One example is the analysis of changes in employability after specific interventions to increase students' self-efficacy, as demonstrated in the Realise 2 Strengths (or R2Strengths) of the LEGACY project¹⁴. The R2Strengths tests if students' changes in employability can be traced after participation in an online careers' guidance tool (*Wilson et al, 2018*, in preparation). R2Strengths uses students' responses to specific questions to identify their strengths, categorised as "realised strengths", "unrealised strengths", "weaknesses" and "learned behaviours". Changes in students' level of career readiness and self-efficacy are analysed using a mixed method design. The rationale is that individuals with a greater degree of self-awareness demonstrate an improved and more accurate articulation, they demonstrate focus, confidence and positivity which can be recognisable during recruitment phases and could be more successful in their career. First results indicate that mainly qualitative methods yield valid and reliable findings to measure changes in students' employability.

¹⁴www.legacy.ac.uk LEGACY stands for Learning and Employability Gain Assessment Community

Other projects have captured student experiences using a combination of qualitative data extracted from participants with some projects linking to alternative data sources facilitating further data analysis. One example for this approach is the pilot project at Ravensbourne to "*Measure Learning Gain from Practice and Work-based Learning Programmes*". Ravensbourne is a University Sector college which a vocational focus on design and digital media industries. With seven collaborative research partners also specialising in creative arts, the learning gain research programme specifically aimed to "*understand the potential to measure the [learning gain] from work-based learning and work preparation activities learning as part of a course of study and the effects of these activities on employability, primarily using existing data held by partner institutions*"¹⁵. It is well established that industrial experience gained in some disciplines/sectors can improve graduates employability by providing a useful insight of industry expectations, operational activity and building up networks. The Ravensbourne project further concludes that longitudinal engagement is necessary, preferable at multiple points through any academic period, furthermore this regular contact with individuals proved beneficial to understanding "career trajectory", "maintain alumni relations" and "work preparation".

Similarly, another strand of the LEGACY project focusses on students' employability gain following international experience gained by students electing to participate in a "work abroad" or "study abroad" scheme. These are compared against a reference group of students that remain "on campus" and their exposure to the diverse student base that are reflected in the student populations. The strand will identify motivations and learning aims, record pre-conceptions of their proposed activity, and perceived benefits of participating in such schemes as well as document any risk averse behaviour. It is anticipated that through this evidence, it will better inform the self-perceived employability gain(s) students achieve from participating in such an experience, improve such schemes for future participating cohorts as well as enhance any pre/post resources. Impact study and analysis of programmes such as ERASMUS (*European Commission, 2016*) have identified there are improvements to a number of perceived skills when participating in such schemes, these are not simply about living in another country, meeting new people, improve/widen career prospects or to learn and improve a foreign language. At a baseline, the scheme is known to improve student's cultural awareness, but the tangible experiences identify skills such as communications, planning and organisational skills, team working, ability to adapt to different situations and problem solving skills.

¹⁵<https://www.ravensbourne.ac.uk/research/research-projects/>

4.2 Methodological considerations

The use of existing instruments to measure the learning gain achieved have yielded some interesting results, for example in *Arum & Roksa's (2011)* ground-breaking work on students' lack of learning gain in the USA. However, consideration must be made on how measurement instruments are implemented, whether they are voluntary or whether they are integrated as part of teaching and learning assessments. Students will want to know the purpose of the measurement and how this would benefit them personally. Whilst the Collegiate Learning Assessment (CLA+) and Grade Point Average (GPA) methods might be appropriate to the needs of the USA education landscape, this does raise questions whether adaptation of this approach would yield an appropriate measurement of learning gain which would be representative of the UK education system. The projects that are testing this approach will need to consider how the calculation of scores reflect the learning gain achieved and what this means to an individual student. It may be advantageous to provide an appropriate explanation of the score achieved, what this means in terms of learning and identification of support services that could improve certain skills, providing a score might be meaningless without appropriate contextual information about the learning gain achieved.

Some pilot projects implement self-assessment tests, these are a recognised form of assessment. Self-assessment can assist in the measurement of meta-cognitive skills and provide an indication of a student's understanding of subject matter. There is flexibility in the range of skills that can be tested and virtual learning environments can assist in ease of assessment. However, there are other considerations such as administration efforts, especially on large-scale deployment and consistency of questions across subjects and different institutions. Early indication in one particular pilot project, which looks specifically at self-efficacy assessment, identifies a link with learning gain and confidence levels being positively correlated where students have been asked to self-assess their performance through multiple choice questions (*Brandriet & Bretz, 2014; Aurah, 2013*).

Some pilot projects have developed new methods bringing together several different dimensions to measure learning gain. The Manchester College is analysing the data of students undertaking higher level courses in a further education setting, combining data from a bespoke learning gain survey and management information system data to determine measurable outcomes of learning gain. Furthermore, it also aims to identify the factors that could affect the measures that have been selected to identify the learning gain. The University of East Anglia¹⁶ are assessing three methods, student marks and GPA, self-efficacy assessment indicators and concept inventories. Some early results in relation to self-efficacy have already been noted within this report. The

¹⁶<https://www.uea.ac.uk/learning-gain/home>

University of Manchester are investigating three ways of measuring the "distance travelled" by students, by analysing a) standardisation of entry/exit qualifications, b) design and validation of new dispositions and critical skills and c) modelling change in learning outcomes. It is anticipated that through quantitative analysis of data, administrative and obtained through longitudinal self-reporting, appropriate assessment will be possible, this is complemented further with interviews and students that have participated in this research allowing for qualitative input to also be analysed accordingly.

Another strand of the LEGACY project, led by the University of Cambridge, has developed a bespoke measurement tool to measure learning gain across disciplines. A three stage approach is used, stage 1 has explored and focussed on students understanding of learning gain through a series of interviews, leading to the (second stage of) development, testing and administration of the tool across a number of consortium partners. The tool is currently applied across two academic years engaging with students who respond to a series of questions that relate to cognitive, meta-cognitive, affective and socio-communicative dimensions; analyses will investigate the relationship between different dimensions and their sub-components to determine the measurement of learning gain. The third stage will involve further meta-analysis of gathered data linked to individual student records. Initial analysis has determined that there is a correlation between deeper patterns of learning and positive attitudes, individuals displaying such behaviours engage more in self-regulatory behaviours. Whilst there will be further student engagement and analysis undertaken to provide a more comprehensive assessment of the different tools and their results, consideration must be paid to factors that are not being measured in any of the new or existing tools which affect a students' experience and consequently their learning gain.

5 Challenges to transfer research findings to policies – a preliminary conclusion

In a market-oriented higher education system such as the UK, the assessment of teaching and learning needs to meet a high level of interest in both academia and policy. The paper discussed how teaching excellence is currently assessed by the Teaching Excellence and Student Outcomes Framework (TEF), and summarised current approaches and methods to measure learning gain which are anticipated to be considered for an implementation of the current practise. The need to measure the "value-added" of tertiary education is justified by a contrast between students' and societal investment in higher education with its employment outcome (*Behle et al., 2015*) and the need to understand more about the potential return to investment. While the need to assess the quality of teaching and learning is fairly uncontested, the way in which policy implements measurements is highly disputed, and the TEF was and is criticised.

The broad variation of institutes, methods employed, and expected results will inevitably yield wide-ranging results which is deliberate and challenging at the same time. It is deliberate as the results represent the breadth of the higher education sector and its diverse body of students. Higher education policy makers will be challenged to integrate proposed methods into existing infrastructure of organisations that have and will make future submissions to the TEF.

As with most pilot programmes, maintaining consistent engagement and retention with participating students requires a specific approach allowing them to remain interested in the project. Whilst some projects have offered incentives to participants, student retention remains a challenging aspect of most pilot projects as many of the projects are longitudinal in their scope. Other methodological drawbacks could be identified with the formation of bespoke or other psychometric tests as, in general, significant financial and resource investment are required for wide implementation.

Despite the ongoing discussions of different ways to measure learning gain and the debate to improve the TEF, the UK example can be seen as a further push towards a larger body of empirical research in the identification of competency in higher education, as demanded by *Zlatkin-Troitschanskaia, Shavelson & Kuhn (2015)*. It will remain problematic, however, to agree on a procedure to measure teaching excellence within the UK, and, on an international level, the identification of learning gain from tertiary study will be even more challenging. In contrast to existing approaches for secondary schools such as the Programme for International Student Assessment (PISA), there is no overall agreement on the level of skills and knowledge that can be expected as an outcome to higher education *per se*. Professional bodies might have established specific levels of skills and knowledge for their particular occupations (e.g. nursing, architecture, engineering), nevertheless, there is no common higher education curriculum that explains the kind and the level of subject-specific and generic skills and knowledge which also results in the lack of transferability of degrees (*Tzanakou & Behle, 2017*). In order to understand and compare different sectors of higher education, agreement on the intellectual outcome of higher education is inevitable.

References

- Arum and Roksa (2011):* Academically Adrift. Limited Learning on College Campuses. University of Chicago Press
- Aurah, C. M. (2013):* The Effects of Self-Efficacy Beliefs and Metacognition on Academic Performance: A Mixed Method Study, *Am. J. Ed. Res*, vol 1, no. 8, p334–343
- Beech, D. (2017):* "Going for Gold. Lessons from the TEF provider submissions." HEPI Report 99

Behle, H., Atfield, G., Elias, P., Gambin, L., Green, A., Hogarth, T., Purcell, K., Tzanakou Ch. and Ch. Warhurst (2015): Reassessing the employment outcomes of higher education. J Huisman and J. Case (eds.) Investigating Higher Education. International perspectives on theory, policy and practice. Oxon and New York, Routledge Press: 114–131

Behle, H. (2016): Graduates in non-graduate occupations. <http://www.hefce.ac.uk/pubs/rereports/year/2016/gradoccup/>

Boud, D. (2018): "Assessment could demonstrate learning gains, but what is required for it to do so?" Higher Education Pedagogies 3(1): 1–3

Brandriet, A. R., Bretz, S. L (2014): Measuring Meta Ignorance through the Lens of Self Confidence: Examining Students' Redox Misconceptions about oxidation Numbers, Charge and Electron Transfer, Che. Educ. Res. Pract., 15, p729–746

Brynin, M. (2013): "Individual Choice and Risk: The Case of Higher Education." Sociology 47(2): 284–300

Clarke, M. (2017): Rethinking graduate employability: the role of capital, individual attributes and context, Studies in Higher Education, DOI: 10.1080/03075079.2017.1294152 (online first)

Department for Education (DfE) (2017): Teaching Excellence and Student Outcomes Framework Specification https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/658490/Teaching_Excellence_and_Student_Outcomes_Framework_Specification.pdf

Douglas, J., Douglas, A. and B. Barnes (2006): "Measuring student satisfaction at a UK university." Quality assurance in education 14(3): 251–267

Elias, P. and Purcell, K. (2004) SOC (HE): A classification of occupations for studying the graduate labour market. <http://www2.warwick.ac.uk/fac/soc/ier/research/completed/7yrs2/rp6.pdf>

European Commission (2016): The Erasmus Impact Study. Regional Analysis http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2016/erasmus-impact_en.pdf

Green, A., de Hoyos, M., Barnes, S-A., Owen, D., Baldauf, B. and Behle, H. (2013): 'Literature Review on Employability, Inclusion and ICT, Report 1: The Concept of Employability, with a Specific Focus on Young People, Older Workers and Migrants', JRC Technical Report (Report EUR 25794 EN), Editors: Institute for Prospective Technological Studies, Joint Research Centre, European Commission, Luxembourg: Publications Office of the European Union, <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=6059>

Harrison, N. and R. Waller (2017): "Success and impact in widening participation policy: What works and how do we know?" Higher Education Policy 30(2): 141–160

Hoareau McGrath, C., Guerin, B. Harte, E. Frearson, M. and C. Manville (2015): Learning Gain in Higher Education, RAND http://www.hefce.ac.uk/media/HEFCE,2014/Content/Pubs/Independentresearch/2015/Learning_gain,in,HE/Learning_gain.pdf (accessed 2017-10-01)

Kandiko Howson, C. (2018): Evaluation of HEFCE's Learning Gain Pilot Projects Year 2. Report to HEFCE by King's College London <https://www.officeforstudents.org.uk/media/1386/evaluation-of-hefce-s-learning-gain-pilot-projects-year-2.pdf> (accessed 2018-04-25)

Naidoo, R. and Jamieson, I. (2005): 'Empowering participants or corroding learning? Towards a research agenda on the impact of student consumerism in higher education', *Journal of Educational Policy*, 20(3), pp. 267–281

Purcell, K., Elias, P., Ellison, R., Atfield, G., Adam, D. and Livano, I. (2008): 'Futuretrack: Applying for Higher Education – Career Choices and Plans, Advance Stage 1 Report', http://www2.warwick.ac.uk/fac/soc/ier/futuretrack/findings/futuretrack_2006_stage1_full_report_2008.pdf (accessed 2017-10-01)

Purcell, K., Elias, P., Atfield, G. Behle, H., R. Ellison and D. Luchinskaya (2013): Transitions into employment, further study and other outcome https://warwick.ac.uk/fac/soc/ier/futuretrack/findings/stage_4_report_final_06_03_2013.pdf

Schleicher, A. (2015): Value-Added. How do you measure whether universities are delivering for their students? HEPI 2015 Annual Lecture. . H. E. P. Institute. HEPI Report 82

Times Higher Education THE (2017): Teaching excellence framework (TEF) results 2017 <https://www.timeshighereducation.com/news/teaching-excellence-framework-tef-results-2017> (accessed 2017-11-20)

Tomlinson, M. (2017): "Student perceptions of themselves as 'consumers' of higher education." *British Journal of Sociology of Education* 38(4): 450–467

Tzanakou, Ch. and H. Behle (2017): The intra-European transferability of graduates' skills gained in the U.K. *Journal of Ethnic and Migration Studies*, Vol 43 (8) pp.1379–1397

Wilson, A., Behle, H. and A. Tassinari (2018, in preparation): Measuring employability gain after R2 Strengths. Research Report

Zlatkin-Troitschanskaia, O., Shavelson, R. J. and Ch. Kuhn (2015): The international state of research on measurement of competency in higher education, *Studies in Higher Education*, 40:3, 393–411

Manuscript submitted: 30-1-2017

Manuscript accepted: 24-04-2018

Address of authors:

Dr. Heike Behle

Sunil Maher

University of Warwick

UK Coventry CV4 7AL

Email: Heike.Behle@warwick.ac.uk

S.Maher.1@warwick.ac.uk

Dr. Heike Behle works at the University of Warwick at the Institute for Employment Research and is currently on a secondment to the LEGACY project. Her main research focuses on the employability gain of graduates.

Sunil Maher is project manager and works for the LEGACY project at the University of Warwick.

Grundlagen der Reputation von Studienstandorten: Bewertungsmaßstäbe für die Qualität von Hochschulstandorten und Studiengängen aus der Sicht von sozialwissenschaftlichen Professorinnen und Professoren

Lorenz Tafertshofer, Evamaria Werner, Bernhard Schmidt-Hertha

Der Beitrag geht der Frage nach, nach welchen Kriterien und auf Basis welcher Informationen Professorinnen und Professoren in den Sozialwissenschaften Studienstandorte und Studiengänge bewerten, und inwieweit diese Reputation ein belastbarer Indikator für Studiengangsqualität sein könnte. Auf Basis einer Onlinebefragung von deutschen Professorinnen und Professoren in fünf sozialwissenschaftlichen Fachbereichen (n=1313) zeigt sich, dass die Reputation eines Standorts oder eines Studiengangs sich vor allem durch die Forschungsreputation der dort lehrenden Kolleginnen und Kollegen sowie die Infrastruktur am Studienort erklärt. Gleichzeitig zeigen sich signifikante Unterschiede zwischen den beteiligten Fächern. Insgesamt scheint die Reputation eines Studienorts oder Studiengangs damit wenig Information über die Qualität der hochschulischen Ausbildung vor Ort zu geben, sondern eher Forschungsindikatoren zu reproduzieren.

1 Einführung

Im Rahmen eines zunehmenden Wettbewerbs in der deutschen Hochschullandschaft (*Trute/Pilniok 2009*) gewinnt auch der Vergleich verschiedener Hochschulstandorte nach unterschiedlichen Bewertungslogiken zunehmend an Bedeutung. Um die Hochschulstandorte jeweils wertend miteinander vergleichen zu können, kommen unterschiedliche Ansätze zum Einsatz. Die Bewertung kann sich dabei auf die gesamte Universität oder nur auf einzelne Fächer beziehen. Ebenso kann nach Aufgabebereichen einer Hochschule wie Lehre und Forschung differenziert werden. Gängig ist beispielweise die Verwendung hochschulstatistischer Kennzahlen wie Drittmittelquoten, die durchschnittliche Studiendauer oder die Anzahl der Promotionen. Auch die Anzahl von Publikationen oder Zitationsquoten werden z.B. in Hochschulrankings oft als Indikatoren für die Forschungsleistung eines Standorts herangezogen. Der Konstruktion vermeintlich geeigneter Indikatoren scheinen dabei keine Grenzen gesetzt zu sein – so wurde im *Academic Ranking of World Universities (ARWU)* die Anzahl an Alumni, die einen Nobelpreis erhalten haben, als Indikator für die Qualität von Studienprogrammen herangezogen (*Kroth/Daniel 2008*). Neben indikatorenbasierten Verfahren

werden auch peer-gestützte Bewertungsverfahren eingesetzt. In verschiedenen Hochschulrankings benennen Wissenschaftlerinnen und Wissenschaftler dabei diejenigen Hochschulstandorte, die aus deren Sicht (in den Bereichen Forschung und Lehre) ein besonders hohes Renommee bzw. eine besonders hohe Reputation besitzen. Solche Reputationsbefragungen sind aus methodischen Gründen (Stichprobengröße, verschiedene Verzerrungseffekte, geringe Reliabilität) nicht unumstritten und die Aussagekraft der Ergebnisse wird teils stark angezweifelt (*Federkeil 2013; Fidler/Parsons 2008; Kroth/Daniel 2008*).

Auch in der vorliegenden Studie wurde im Rahmen der Online-Befragung je fachspezifisch erfragt, welche Hochschulstandorte die befragten sozialwissenschaftlichen Professorinnen und Professoren als überdurchschnittlich gut in Forschung sowie in Lehre erachten. Zudem wurden die Professorinnen und Professoren gebeten, je vier Studiengänge zu benennen, die sie als überdurchschnittlich gut hinsichtlich ihrer Qualität ansehen würden. Neben dem Renommee von Hochschulstandorten in den Bereichen Forschung und Lehre wurde der Blick also vertiefend auf die spezifischen Studiengänge gerichtet. Für die forschungsleitenden Fragestellungen war dabei weniger von Interesse, welche Standorte und Studiengänge tatsächlich von den Befragten benannt wurden, als vielmehr, welche Bewertungsmaßstäbe den jeweiligen Nennungen zugrunde lagen. Es soll also untersucht werden, auf welche Kriterien sich die positiven Bewertungen der Professorinnen und Professoren stützen, d.h. worauf eine hohe Reputation eines Standorts in der Forschung oder in der Lehre bzw. eine hohe Reputation von Studiengängen beruht. Dabei sind Unterschiede zwischen verschiedenen Hochschularten vor dem Hintergrund eines in Hochschulen für angewandte Wissenschaften nicht expliziten Forschungsauftrags der Hochschullehrenden zu diskutieren.

Hierzu werden zunächst einige allgemeine Überlegungen zum Thema Reputation in der Wissenschaft vorgestellt, um dann näher auf Material und Methode einzugehen, bevor im nächsten Schritt die Ergebnisse der empirischen Untersuchung präsentiert werden. Dabei sollen auch fachspezifische Unterschiede bei den herangezogenen Bewertungsmaßstäben zwischen Befragten der Fächer Erziehungswissenschaft, Politikwissenschaft, Psychologie, Sozialwesen, Soziologie sowie den Wirtschaftswissenschaften herausgearbeitet werden. Im abschließenden Fazit soll mit Rückgriff auf die Befragungsergebnisse dann unter anderem diskutiert werden, wie brauchbar die Reputation eines Hochschulstandorts als ein Indikator für die Qualität von Forschung oder Lehre – als welcher die Reputation in Hochschulrankings ja gerne herangezogen wird – überhaupt sein kann.

2 Reputation in der Wissenschaft

Der Gewinn von Reputation innerhalb der eigenen Scientific Community wird als einer der zentralen motivational wirksam werdenden Anreize für das Engagement von Professorinnen und Professoren im wissenschaftlichen Tätigkeitsfeld gesehen (vgl. *Enders 1998; Honecker 2011; Kühl 2015; Minssen et al. 2003*). Dabei kann angenommen werden, dass dies aufgrund der höheren Sichtbarkeit des eigenen Forschungsoutputs für andere Fachkolleginnen und Fachkollegen (z. B. in Form von Publikationen oder Vorträgen auf Tagungen) für den Bereich Forschung in höherem Maße zutrifft als für den Bereich Lehre. Zwar kann auch die Lehrleistung einzelner Wissenschaftlerinnen und Wissenschaftler nach außen sichtbar gemacht werden, indem etwa öffentlichkeitswirksam Lehrpreise für herausragende Leistungen vergeben werden. Dies geschieht allerdings noch in relativ geringem Ausmaß, auch wenn in den letzten Jahren Bemühungen zur Aufwertung der Lehre zu erkennen sind (vgl. *Wissenschaftsrat 2008*). Studentische Lehrvaluationen stellen eine weitere Möglichkeit dar, die Lehrleistung nach außen sichtbar zu machen. Außer im Rahmen von Berufungsverfahren oder zum universitätsinternen Benchmarking werden personenbezogene Lehrvaluationen aber meist nicht offengelegt.¹ Die Aussagekraft von studentischen Lehrvaluationen als Indikator für die Lehrqualität ist zudem nicht unumstritten und die Ergebnisse können teils durch äußere Einflussfaktoren verzerrt sein (*Schmidt/Weishaupt 2008; Wolbring 2013*).

Neben der Forschungsreputation und der Lehrreputation können noch andere, externe Stakeholdergruppen wie z. B. wirtschaftliche Entscheidungsträger (Drittmitter), Medien oder privatwirtschaftliche Kooperationspartner im Rahmen des Reputationsmanagements eine Rolle spielen (vgl. *Armbruster/Lederbogen 2010*). Der Reputation innerhalb des eigenen akademischen Feldes wird aber eine deutlich höhere Bedeutung zugesprochen als der Reputation im außerwissenschaftlichen Bereich (*Honecker 2011*). Anzunehmen ist nach den obigen Ausführungen zudem, dass die Reputation von Wissenschaftlerinnen und Wissenschaftlern innerhalb des eigenen akademischen Feldes primär deren Forschungsreputation entspricht bzw. auf dieser gründet (vgl. *Armbruster/Lederbogen 2010*).

Aber nicht nur für die Wissenschaftlerinnen und Wissenschaftler, sondern auch für die einzelnen Hochschulstandorte spielt die Reputation innerhalb und außerhalb des wissenschaftlichen Feldes eine wichtige Rolle: Eine hohe Reputation eines Standorts erhöht die Attraktivität eines Standorts für Wissenschaftlerinnen und Wissenschaftler, für (potentielle) Studierende und für privatwirtschaftliche Kooperationspartner. Die

¹In Hochschulrankings, wie z. B. dem CHE-Ranking, werden Studierendenbewertungen allerdings sehr wohl nach außen sichtbar gemacht. Diese Bewertungen beziehen sich aber nicht auf einzelne Lehrpersonen, sondern auf verschiedene Aspekte (z. B. Studierbarkeit insgesamt, Betreuung durch Lehrende, etc.) des gesamten Lehrangebotes des betroffenen Faches eines Standorts.

Kriterien, die für die verschiedenen Stakeholder hierbei relevant werden, können dabei ganz unterschiedlicher Natur sein. Während für Studierende vor allem die Lehrsituation an der Hochschule, z.B. vermittelt über Erfahrungsberichte anderer Absolventen, von Interesse sein dürfte, sind für die privatwirtschaftliche oder öffentliche Auftragsforschung das Forschungsprofil und der bisherige Forschungoutput von höherer Relevanz. Die Reputation eines Hochschulstandorts kann damit als „aggregiertes Bündel“ (Honecker 2011, S. 18) von Fremdwahrnehmungen der verschiedenen Stakeholder verstanden werden. In der vorliegenden Studie wird mit den sozialwissenschaftlichen Professorinnen und Professoren eine der wichtigsten Stakeholdergruppen herausgegriffen und untersucht, welche Bewertungsmaßstäbe für die Einstufung eines Hochschulstandorts als in Lehre oder Forschung herausragenden Standort herangezogen werden. Damit wird also nur ein Teil des oben genannten „Reputationsbündels“ analysiert, während die Wahrnehmungen anderer Stakeholdergruppen unberücksichtigt bleiben müssen.

Hinsichtlich der Reputation eines Hochschulstandorts aus Sicht von Professorinnen und Professoren ist dabei von besonderem Interesse, welche Kriterien zur Beurteilung der – von außen wenig sichtbaren – Lehre herangezogen werden (können) bzw. auf welchen Kriterien die Lehrreputation eines Standorts beruht. Hierzu werden neben den Bewertungsmaßstäben für die Lehrreputation eines Standorts auch die Bewertungsmaßstäbe zur Beurteilung herausragender Studiengänge in den Blick genommen. Zudem ist sowohl für den Bereich Lehre als auch für den Bereich Forschung anzunehmen, dass sich zwischen den untersuchten Fächern (Erziehungswissenschaft, Politikwissenschaft, Psychologie, Sozialwesen, Soziologie und Wirtschaftswissenschaften) aufgrund unterschiedlicher Fachkulturen auch Unterschiede bezüglich der Relevanz verschiedener Bewertungsmaßstäbe ergeben. Daher soll im Folgenden auch überprüft werden, inwieweit sich signifikante Fachunterschiede nachweisen lassen.

3 Material und Methode

Die im Folgenden verwendeten Daten stammen aus einer Online-Befragung von Professorinnen und Professoren, die im Rahmen des Projekts „QualRep: Qualität von sozialwissenschaftlichen Studiengängen und Reputation von Studienstandorten aus Sicht von Professorinnen und Professoren“ realisiert wurde. Das Projekt wurde im Rahmen des BMBF-Programms „Leistungsbewertung in der Wissenschaft“ gefördert (Förderkennzeichen: 01PY13005) und umfasste neben der hier im Zentrum stehenden Online-Befragung auch weitere quantitative sowie qualitative Forschungszugänge (vgl. Schmidt-Hertha/Thalhammer/Müller 2015).

3.1 Stichprobe

Für die Befragung wurden die E-Mail-Adressen von Hochschullehrenden aus sozialwissenschaftlichen Fachbereichen (nur staatliche Hochschulen)² in ganz Deutschland ermittelt und der Fragebogen an diese versendet. Zur Ermittlung der Hochschulen wurden die Listen des Statistischen Bundesamtes herangezogen. Die E-Mail-Adressen der Hochschullehrenden wurden mithilfe des Hochschullehrerverzeichnisses (Jahr 2013) sowie einer manuellen Nachrecherche festgestellt. Der Link zur Onlinebefragung wurde an 3786 Hochschullehrende versendet, schlussendlich konnten 1313 Fragebögen verwertet werden, was einer Rücklaufquote von 34 Prozent entspricht. 1,5 Prozent der Befragten waren als Juniorprofessorin oder Juniorprofessor (Besoldungsgruppe W1) beschäftigt, 55,4 Prozent als Professorin oder Professor in der Besoldungsgruppe W2 und W3 und 24,8 Prozent als Professorin oder Professor mit Besoldungsgruppe C2–C4. Zusätzlich nahmen auch einige außerplanmäßige Professorinnen oder Professoren (1,8 Prozent) und Hochschullehrende mit anderem Status (3,0 Prozent) an der Befragung teil. 13,4 Prozent machten keine Angaben zu ihrem beruflichen Status. Die Stichprobe verteilt sich wie folgt auf die verschiedenen Hochschularten: 41,1 Prozent Universität, 39,6 Prozent Hochschule für angewandte Wissenschaften, 1,1 Prozent Pädagogische Hochschule und 18,3 Prozent ohne Angabe. Für die Analysen besonders relevant war die disziplinäre Zuordnung der Befragten, die sich zu 11,6 Prozent der Erziehungswissenschaft/Sonderpädagogik, zu 36,1 Prozent den Wirtschaftswissenschaften, zu 7,3 Prozent der Politikwissenschaft, zu 8,5 Prozent der Soziologie, zu 13,6 Prozent dem Bereich Sozialwesen und zu der Psychologie zuordneten. 167 Befragte (12,7 Prozent) machten zu ihrer fachlichen Zugehörigkeit keine Angaben und mussten deshalb bei den fachspezifischen Analysen ausgeschlossen werden.

Damit ist die Stichprobenstruktur hinsichtlich der genannten Merkmale der Gesamtpopulation zumindest ähnlich. Während die Verteilung auf Hochschularten und die Geschlechterverteilung in der Stichprobe nur moderat von den Daten der Hochschulstatistik abweicht, sind C-Professuren und vor allem die Wirtschaftswissenschaften als mit Abstand größtes Fach in der Stichprobe deutlich unterrepräsentiert, zu Gunsten der W-Professuren bzw. aller anderen Fächer (siehe Tabelle 1). Dabei ist zu berücksichtigen, dass insbesondere bei Fragen zur Person oder beruflicher Situation – die bewusst nicht als Pflichtangaben programmiert waren – häufig keine Angaben gemacht wurden. Aufgrund der Rückmeldungen per E-Mail oder in den offenen Fragen wurde verschiedentlich darauf hingewiesen, dass die Kombination der unterschiedlichen Angaben unter Umständen Rückschlüsse auf eine konkrete Person erlaubt und deshalb

²Nicht berücksichtigt wurden Hochschulen in privater oder kirchlicher Trägerschaft, Kunst- und Musikhochschulen sowie Verwaltungshochschulen, da diese aufgrund der Finanzierungsstruktur bzw. der inhaltlichen Ausrichtung Sonderfälle darstellen.

einige Angaben nicht gemacht wurden. Vor diesem Hintergrund ist nicht davon auszugehen, dass die fehlenden Angaben hier zu einer systematischen Verzerrung führten.

Tabelle 1: Stichprobenstruktur in Abgleich mit Daten der amtlichen Statistik

	Stichprobe*	Professuren in Rechts-, Wirtschafts- und Sozialwissenschaften
W2/W3-Professuren	67,8 %	59,1 %
C2/C3/C4-Professuren	30,4 %	36,9 %
Juniorprofessuren	1,8 %	4,0 %
Anteil Professorinnen	30,0 %	29,1 %
	Stichprobe*	Professuren in Deutschland
Universitäten	50,2 %	54,9 %
Hochschulen für angewandte Wissenschaften	48,4 %	44,2 %
Pädagogische Hochschulen	1,3 %	0,8 %
Erziehungswissenschaft	13,3 %	10,1 %
Wirtschaftswissenschaften	41,4 %	58,8 %
Politikwissenschaft	8,4 %	3,5 %
Soziologie/Sozialwissenschaft	9,7 %	6,7 %
Sozialwesen	15,6 %	12,6 %
Psychologie	11,7 %	8,3 %

*Anteil an den gültigen Antworten

Quelle: Statistisches Bundesamt 2017 und eigene Berechnungen

3.2 Befragungsinstrument

Mit Hilfe der Online-Befragung sollte herausgefunden werden, welche Instrumente und Verfahren der Leistungsbewertung und Leistungsmessung an Hochschulen präsent sind und welche die Befragten als Anreize für ein höheres Engagement in den Bereichen Forschung und Lehre sehen. Hierbei wurden die Bereiche Forschung und Lehre, welche beide zum Arbeitsalltag von Hochschullehrenden gehören, getrennt voneinander betrachtet. In der Lehre sind Formen der Leistungsbewertung und -messung beispielsweise vergebene Lehrpreise, Absolventenbefragungen, externe Lehrevaluationen, Lehrevaluation durch Studierende, Lehrberichte, Akkreditierung oder Hochschulrankings, die teilweise auch Einfluss auf Karrierechancen haben können.

Ebenso gibt es in der Forschung bestimmte Verfahren und Methoden der Leistungsmessung und Leistungsbewertung sowie den verschiedenen Anreizen, die an Hochschulen Verwendung finden. Eine Auswahl davon, die wir im Fragebogen vorgegeben haben, umfasst (eingeworbene) Drittmittel, Vortragstätigkeiten, Funktionen in Fach-

gesellschaften, Kooperation mit Unternehmen, Publikationen in einer referierten Zeitschrift, Anzahl der Publikationen, (betreute) Promotionen, (betreute) Habilitationen, Anzahl Zitationen/Zitationskennzahlen, internationale Kooperationen, Forschungstransfer und Hochschulrankings. Weiter kann die Liste ergänzt werden mit Deputatsreduktionen, zusätzliche Mittel für die Forschungsarbeit, Anerkennung durch die Hochschulleitung, öffentliche Anerkennung der Leistungen, Forschungspreise, Forschungsreputation, leistungsorientierte Mittelvergabe oder Karrierechancen. Die Bedeutung der jeweiligen Indikatoren für hochschulinterne Steuerungsprozesse variiert sowohl zwischen verschiedenen Fakultäten und Fachbereichen als auch zwischen verschiedenen Hochschulen und diese sind auch in den verschiedenen Scientific Communities unterschiedlich anerkannt. Intention der vorliegenden Befragung war es auch, zu erfahren, inwiefern bestimmte Leistungsbewertungsverfahren und Anreize von den Befragten als geeignet und sinnvoll empfunden werden. Die Ergebnisse hierzu können im Rahmen dieses Beitrags allerdings nicht vertieft werden.

Der Fragebogen wurde in verschiedene thematische Einheiten unterteilt, die jeweils aus Gruppen einzelner Items bestanden. Für diese Items wurde – sofern nicht anders angegeben – jeweils mit Likertskalen gearbeitet von 1 (sehr geeignet/stimme voll und ganz zu) bis 6 (überhaupt nicht geeignet/stimme überhaupt nicht zu). Die Befragten wurden im Fragebogen gebeten, sich dazu zu äußern, welche Studiengänge und Studienstandorte sie in ihrem Fachbereich jeweils als herausragend im Bereich Lehre und Forschung empfinden. Im Anschluss wurde gefragt, welche Referenzkategorien von den Befragten für ihre Einschätzung herangezogen wurden. Mit diesem Vorgehen wurde versucht die Bewertung von Indikatoren und anderen Informationsquellen an einen konkreten Entscheidungs- und Bewertungsprozess zu knüpfen, die einer Reputationserhebung in Hochschulrankings ähnelt.

4 Bewertungsmaßstäbe für die Qualität von Hochschulstandorten und Studiengängen

Um zu ermitteln, welche Kriterien von Professorinnen und Professoren zur Beurteilung von in Forschung und Lehre herausragenden sozialwissenschaftlichen Hochschulstandorten herangezogen werden, wurden diese im Rahmen der Online-Befragung gebeten, jeweils vier Standorte ihres Fachs zu benennen, welche als überdurchschnittlich gut in der Forschung und in der Lehre angesehen werden. Die Abfrage wurde dabei jeweils getrennt für den Bereich Forschung und den Bereich Lehre durchgeführt. Im Anschluss daran wurde wieder jeweils separat für Forschung und Lehre erfragt, auf welchen Kriterien die vorangegangenen Einschätzungen beruhen. Die Abfrage erfolgte in Form von halboffenen Fragen mit vorgegebenen Antwortmöglichkeiten und einer „Sonstige“-Kategorie mit Freitextfeld. Analog wurde bei der Beurteilung der Studiengänge und der Abfrage der hierfür relevanten Bewertungsmaßstäbe verfahren.

Die prozentualen Antworthäufigkeiten zu den vorgegebenen Kriterien sind in Tabelle 2 für den Bereich Lehre, in Abbildung 2 für die Studiengänge und in Tabelle 3 für den Bereich Forschung dargestellt.

Ein kurzer Hinweis zu den differenten Fallzahlen: die Fragen zu den folgend abgebildeten Ergebnissen mussten nicht zwingend beantwortet werden um den Fragebogen weiter führen zu können.

Tabelle 2: Antworthäufigkeiten herangezogene Bewertungsmaßstäbe Lehre (n = 653)

	Eigene Erfahrungen	Reputation von Kolleginnen & Kollegen	Hochschulrankings	Auswahlgespräche mit Absolventinnen & Absolventen	Teilnahme an Evaluationen/ Akkreditierungen als Gutachter*in	An den Standort vergebene Lehrpreise	Materielle Infrastruktur (z.B. Bibliothek)	Personelle Infrastruktur des Standorts (z. B. Professor*innen, Breite des Faches)
Erziehungswissenschaft (n=108)	28,7 %	54,6 %	3,7 %	13,9 %	11,1 %	0 %	11,1 %	44,4 %
Wirtschaftswissenschaft Uni (n=110)	35,5 %	58,2 %	20,9 %	22,7 %	6,4 %	0 %	5,5 %	40,9 %
Wirtschaftswissenschaft FH (n=222)	42,8 %	58,1 %	22,1 %	14,9 %	8,6 %	3,6 %	9,9 %	27,0 %
Soziologie (n=80)	35,0 %	55,0 %	5,0 %	27,5 %	1,3 %	0 %	8,8 %	32,5 %
Politikwissenschaft (n=66)	43,9 %	81,8 %	18,2 %	24,2 %	9,1 %	0 %	16,7 %	62,1 %
Psychologie (n=67)	40,3 %	44,8 %	10,4 %	11,9 %	4,5 %	3,0 %	13,4 %	35,8 %
Sozialwesen (n=142)	31,0 %	60,6 %	3,5 %	7,7 %	8,5 %	2,1 %	5,6 %	32,4 %

Quelle: Eigene Berechnungen

Für den Bereich Lehre zeigt sich, dass insbesondere die Reputation von Kolleginnen und Kollegen ein wichtiges Kriterium zur Beurteilung der Lehre anderer Standorte darstellt. Auch die personelle Infrastruktur wird von etwas mehr als einem Drittel der Befragten genannt, wobei hier von inhaltlichen Überschneidungen zum Item *Reputation von Kollegen* ausgegangen werden muss. Daneben spielen auch eigene Erfahrungen (selbst an dem Standort tätig gewesen) mit einer Antworthäufigkeit von knapp 37 Prozent eine wichtige Rolle als Bewertungsmaßstab. Auswahlgespräche mit Absolventinnen und Absolventen sowie Hochschulrankings sind mit einem Anteil von ca. 17 und 13 Prozent von geringerer Bedeutung, wobei sich allerdings teils deutliche Unterschiede zwischen den einzelnen Fächern ergeben.

Mittels einfaktorieller Varianzanalysen konnten bei beiden Items jeweils signifikante Fachunterschiede nachgewiesen werden³: Hinsichtlich der Hochschulrankings ($F(5, 795) = 9.51, p < .001$) zeigt sich, dass diese im Fach Wirtschaftswissenschaften mit einem Anteil von etwa 21 Prozent signifikant häufiger als Bewertungsmaßstab für die Lehre anderer Standorte dienen als in der Erziehungswissenschaft (4 Prozent), der Sozialen Arbeit/Sozialwesen (4 Prozent) und der Soziologie (5 Prozent). Dies ist sicherlich dem höheren Stellenwert von Rankings in den Wirtschaftswissenschaften (man denke zum Beispiel an das Handelsblattranking) bzw. der tendenziell positiveren Einstellung⁴ gegenüber diesen geschuldet. Auswahlgespräche mit Absolventinnen und Absolventen ($F(5, 795) = 3.79, p = .002$) werden dagegen insbesondere in der Soziologie (28 Prozent) signifikant öfter genannt als im Fach Sozialwesen (8 Prozent), was auch mit unterschiedlichen Auswahlkulturen an Universitäten und Hochschulen für angewandte Wissenschaften zusammenhängen könnte. Auch die Antworthäufigkeiten zu den Kriterien *Reputation von Kolleginnen und Kollegen* ($F(5, 795) = 4.25, p = .001$) sowie *Personelle Infrastruktur des Standorts* ($F(5, 795) = 5.24, p < .001$) unterscheiden sich signifikant zwischen einzelnen Fächern: Die Reputation von Kolleginnen und Kollegen wird von Befragten aus der Politikwissenschaft mit einem Anteil an Nennungen von gut 80 Prozent signifikant öfter als Bewertungsmaßstab herangezogen als von Befragten der Fächer Erziehungswissenschaft (55 Prozent), Psychologie (45 Prozent), Soziologie (55 Prozent) und den Wirtschaftswissenschaften (58 Prozent). Ähnlich verhält es sich bei der personellen Infrastruktur. Diese wird im Fach Politikwissenschaft von 61 Prozent der Befragten als Bewertungsmaßstab genannt, während der Anteil an Nennungen in den Fächern Psychologie (36 Prozent), Soziologie (33 Prozent), Sozialwesen (32 Prozent) sowie Wirtschaftswissenschaften (31 Prozent) signifikant niedriger ausfällt. Die Erziehungswissenschaft liegt mit einem Wert von 44 Prozent dazwischen (siehe auch Tabelle 2). Dass personelle Aspekte bzw. die Reputation von Kolleginnen und Kollegen für die Bewertung der Lehre anderer Standorte in der Politikwissenschaft eine deutlich größere Rolle spielen als in den anderen Fächern, hängt sicherlich auch damit zusammen, dass die Politikwissenschaft mit nur 373 Professorinnen und Professoren (Stand 2014) im Vergleich das kleinste Fach darstellt (*Statistisches Bundesamt 2015*).⁵ In einer kleineren Fachcommunity sind die einzelnen Professorinnen und

³Aufgrund fehlender Varianzhomogenität wurden zur Überprüfung der Ergebnisse des F-Tests auch der robuste Welch-Test sowie der Brown-Forsythe-Test herangezogen. Die Ergebnisse konnte dabei durchgehend bestätigt werden. Berichtet werden hier die Werte des F-Tests. Zur Ermittlung der signifikanten Gruppenunterschiede (welches Fach weicht signifikant von den anderen Fächern ab) wurde entsprechend der fehlenden Varianzhomogenität neben der Scheffé-Prozedur zusätzlich der Games-Howell-Test zum Post-Hoc-Mehrfachvergleich herangezogen.

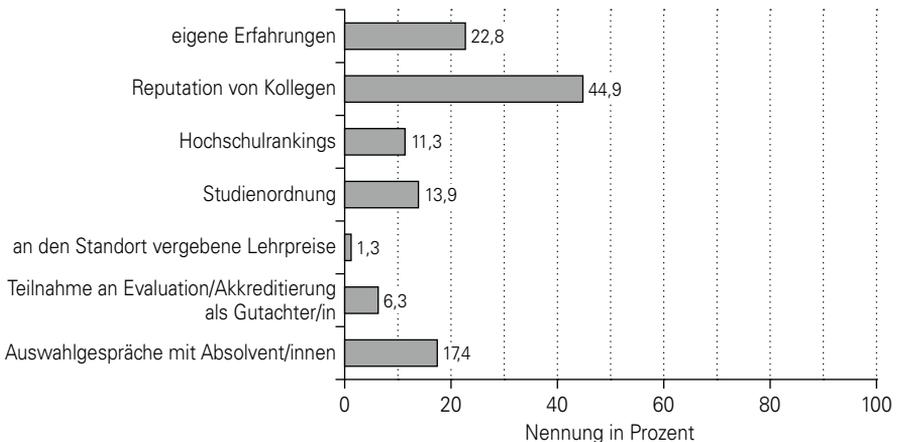
⁴Im Rahmen des Projekts wurde in der Online-Befragung auch die Einstellung gegenüber Rankings erhoben. Hierbei zeigte sich eine leicht positivere Einstellung gegenüber Rankings bei den Professorinnen und Professoren der Wirtschaftswissenschaften (erhoben mittels Ratingskalen), auch wenn die Werte insgesamt eine immer noch eher kritische Haltung gegenüber Rankings offenbarten.

⁵Im Vergleich dazu sind beispielsweise in der Psychologie 891 und den Wirtschaftswissenschaften sogar 6297 Professorinnen und Professoren beschäftigt (*Statistisches Bundesamt 2015*).

Professoren sichtbar als in einer sehr großen, weniger überschaubaren Community und können entsprechend leichter als Bewertungskriterium herangezogen werden.

Für die spezifische Beurteilung der Studiengänge ergibt sich ein recht ähnliches Bild wie bei der Beurteilung der in der Lehre generell herausragenden Standorte. Dies ist wenig verwunderlich, da die Qualität von Studiengängen einen Teil der generellen Lehrqualität eines Standorts darstellt und somit auch große Überschneidungen hinsichtlich der Relevanz verschiedener Bewertungsmaßstäbe zu erwarten waren. Die Reputation von Kolleginnen und Kollegen und eigene Erfahrungen werden zwar seltener, aber wie auch im Bereich Lehre dennoch am häufigsten genannt. Die geringere Anzahl an Nennungen beim Item *Eigene Erfahrungen* lässt sich schlicht damit erklären, dass eigene Erfahrungen mit spezifischen Studiengängen seltener gemacht werden (können) als Erfahrungen mit der Lehre an einem anderen Hochschulstandort generell. Ebenso ist einer Professorin oder einem Professor vermutlich eher bekannt, an welchem Hochschulstandort eine hoch reputierliche Kollegin bzw. ein hoch reputierlicher Kollege arbeitet, als in welchem spezifischen Studiengang diese bzw. dieser aktuell tätig ist.

Abbildung 1: Antworthäufigkeiten herangezogene Bewertungsmaßstäbe Studiengang
(n = 639)



Quelle: Eigene Berechnungen

Für die Bewertung von Studiengängen ist die Gewichtung der herangezogenen Kriterien sehr ähnlich. Zudem beruht die Beurteilung der Studiengänge bei knapp 14 Prozent der Befragten auf den Studienordnungen der jeweiligen Studiengänge.⁶

⁶Das Item „Studienordnung“ war bei der generellen Bewertung der Lehre der Hochschulstandorte im Fragebogen nicht als Bewertungsmaßstab vorgegeben.

Betrachtet man die Unterschiede der Antworthäufigkeiten zwischen den Fächern ergeben sich ebenfalls übereinstimmend mit den vorangegangenen Analysen ähnliche signifikante Differenzen. Zur Beurteilung der Lehre eines Standorts generell und zur Beurteilung von Studiengängen eines Standorts im Speziellen werden also von den sozialwissenschaftlichen Professorinnen und Professoren überwiegend dieselben Bewertungsmaßstäbe herangezogen.

Für den Bereich Forschung zeigt sich zunächst ebenfalls ein ähnliches Bild wie für den Bereich Lehre. Auch hier ist die Reputation von Kolleginnen und Kollegen das am häufigsten genannte Kriterium, wobei der Anteil der Nennungen mit über 69 Prozent noch deutlich höher ausfällt als im Bereich Lehre (siehe Tabelle 3).

Tabelle 3: Antworthäufigkeiten herangezogene Bewertungsmaßstäbe Forschung (n = 589)

	Eigene Erfahrungen	Reputation von Kolleginnen & Kollegen	Hochschulrankings	Teilnahme an Evaluationen/ Akkreditierungen als Gutachter*in	Wissenschaftlicher Nachwuchs (Promotion/Habilitation)	An den Standort vergebene Forschungspreise	Sichtbarkeit von Forschungsprojekten	Drittmittelaufkommen	Materielle Infrastruktur (z. B. Bibliothek)	Personelle Infrastruktur des Standorts (z. B. Professuren, Breite des Fachs)	Publikationen
Erziehungswissenschaft (n=108)	20,6%	71,1%	9,3%	9,3%	33%	0%	60,8%	30,9%	8,2%	33%	53,6%
Wirtschaftswissenschaft Uni (n=110)	34,0%	82,5%	30,9%	4,1%	38,1%	13,4%	45,4%	11,3%	4,1%	37,1%	70,1%
Wirtschaftswissenschaft FH (n=222)	26,6%	61,3%	18,1%	4,0%	18,1%	11,1%	39,2%	11,6%	7,5%	23,6%	46,7%
Soziologie (n=80)	34,7%	70,8%	8,3%	5,6%	36,1%	2,8%	56,9%	19,4%	5,6%	40,3%	59,7%
Politikwissenschaft (n=66)	43,8%	84,4%	21,9%	7,8%	37,5%	14,1%	76,6%	42,2%	10,9%	51,6%	71,9%
Psychologie (n=67)	26,2%	73,8%	18%	1,6%	36,1%	4,9%	59%	36,1%	3,3%	29,5%	54,1%
Sozialwesen (n=142)	21,3%	59,8%	4,7%	3,9%	22,8%	3,9%	47,2%	13,4%	1,6%	15,7%	40,9%

Quelle: Eigene Berechnungen

Daneben sind Publikationen sowie die Sichtbarkeit von Forschungsprojekten mit einem Anteil an Nennungen von je über 50 Prozent diejenigen Bewertungsmaßstäbe, die am häufigsten zur Beurteilung der Forschung herangezogen wurden. Es folgen der wissenschaftliche Nachwuchs, die personelle Infrastruktur sowie eigene Erfahrungen, welche von jeweils etwas weniger als einem Drittel der Befragten genannt wurden. Der auch in Rankings und verschiedenen Evaluationsverfahren oft verwendete Forschungsindikator des Drittmittelaufkommens diente nur gut 20 Prozent der Befragten als Kriterium. Hochschulrankings selber sind für immerhin knapp 16 Prozent ein Bewertungsmaßstab für die Forschungsleistung anderer Standorte. Die übrigen vorgegebenen Kriterien wurden insgesamt nur von einem kleinen Teil der Befragten zur Bewertung herangezogen.

Wie bei den vorangegangenen Analysen ergeben sich auch für den Bereich Forschung signifikante Fachunterschiede. Wieder wird die Reputation von Kollegen ($F(5, 714) = 2.34, p = .040$) von den Befragten aus der Politikwissenschaft am häufigsten als Bewertungsmaßstab herangezogen (84 Prozent). Signifikant geringer fällt der Anteil an Nennungen im Fach Sozialwesen (61 Prozent) und den Wirtschaftswissenschaften (68 Prozent) aus. In den Wirtschaftswissenschaften zeigt sich allerdings bei diesem Kriterium ein deutlicher Unterschied zwischen Professorinnen und Professoren an Universitäten und solchen an Hochschulen für angewandte Wissenschaften. Letztere ziehen die Reputation von Kolleginnen und Kollegen als Bewertungsmaßstab deutlich seltener heran (61 Prozent) als wirtschaftswissenschaftliche Universitätsprofessorinnen und -professoren (83 Prozent). Der signifikant niedrigere Wert in den Wirtschaftswissenschaften ist also auf die Befragten mit Professuren an Hochschulen für angewandte Wissenschaften zurückzuführen, was auch für den Bereich Sozialwesen – der überwiegend an Hochschulen für angewandte Wissenschaften angeboten wird – gelten dürfte. Da Professorinnen und Professoren an Hochschulen für angewandte Wissenschaften keinen expliziten Forschungsauftrag haben, spielt auch das Thema Reputation in diesem Hochschulsegment eine deutlich geringere Rolle.

Ähnliches wie bei der Reputation von Kolleginnen und Kollegen gilt für das Kriterium der personellen Infrastruktur ($F(5, 714) = 6.37, p < .001$): Der Anteil an Nennungen im Fach Politikwissenschaft liegt hier bei 52 Prozent, während dieser in den Fächern Sozialwesen und Wirtschaftswissenschaften nur 16 bzw. 28 Prozent beträgt. Für die wirtschaftswissenschaftlichen Universitätsprofessuren ergibt sich wieder ein deutlich höherer Anteilswert von 37 Prozent, während Professorinnen und Professoren an Hochschulen für angewandte Wissenschaften nur zu knapp 24 Prozent die personelle Infrastruktur als Kriterium heranziehen. Hochschulrankings ($F(5, 714) = 5.69, p < .001$) werden mit einem Anteil von knapp 22 Prozent erneut von den Befragten aus den Wirtschaftswissenschaften am häufigsten genannt. Signifikant niedriger fällt der Anteil

an Nennungen dagegen bei den Fächern Erziehungswissenschaft (9 Prozent), Soziologie (8 Prozent) und Sozialwesen (5 Prozent) aus. Als Bewertungsmaßstab spielen Hochschulrankings also vor allem in den Wirtschaftswissenschaften eine Rolle, wobei diese wiederum von Universitätsprofessorinnen und -professoren deutlich häufiger (31 Prozent) als Bewertungskriterium genutzt werden als von Professorinnen und Professoren an Hochschulen für angewandte Wissenschaften (18 Prozent). Dies könnte darauf zurückzuführen sein, dass Hochschulen für angewandte Wissenschaften gerade in internationalen Rankings oft nicht berücksichtigt werden und somit für die dort tätigen Professorinnen und Professoren im beruflichen Alltag eher von geringerer Bedeutung sind.

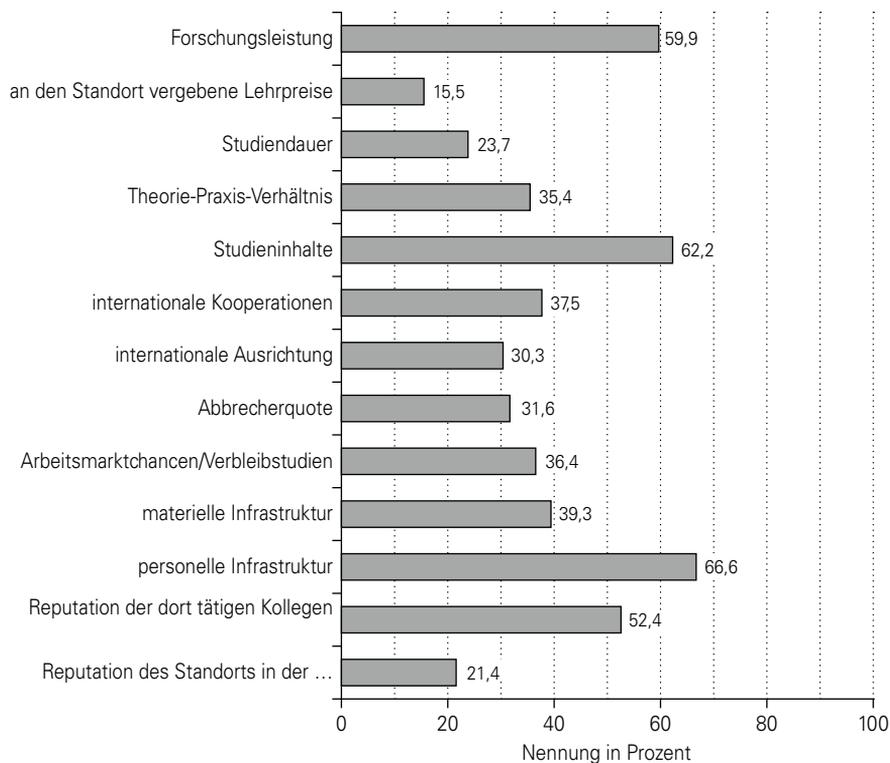
Bei dem klassischen Forschungsindikator des Drittmittelaufkommens ($F(5, 714) = 11.23$, $p < .001$) zeigen sich ebenfalls signifikante Unterschiede zwischen einzelnen Fächern. Das Drittmittelaufkommen wurde vor allem von Befragten der Psychologie (36 Prozent), der Politikwissenschaft (42 Prozent) und – in etwas geringerem Ausmaß – von Befragten der Erziehungswissenschaft (31 Prozent) als Bewertungsmaßstab herangezogen. Das „Schlusslicht“ bilden die Wirtschaftswissenschaften (12 Prozent) und das Fach Sozialwesen (14 Prozent), während sich die Soziologie mit einem Anteil von 19 Prozent im mittleren Bereich bewegt. Aufgrund des – gemessen am gesamten Drittmittelvolumen (vgl. *Statistisches Bundesamt 2016*) – geringeren Stellenwerts von Drittmittelforschung im Fach Sozialwesen verwundert dieses Ergebnis nicht. Auch für das Fach Wirtschaftswissenschaften wäre zu vermuten, dass das Drittmittelaufkommen als Bewertungsmaßstab vor allem für die Professorinnen und Professoren an Hochschulen für angewandte Wissenschaften von geringerer Bedeutung ist. Dies ist aber nicht der Fall – das Drittmittelaufkommen wird von je etwa 11 (Universitäten) bzw. 12 Prozent (Hochschulen für angewandte Wissenschaften) der Befragten aus den Wirtschaftswissenschaften als Kriterium herangezogen. Möglicherweise ist die Drittmittellandschaft aufgrund der Größe des Fachs in den Wirtschaftswissenschaften weniger überschaubar als in anderen Fächern, so dass dem Drittmittelaufkommen als Bewertungsmaßstab eine nur untergeordnete Rolle zukommt.

Die Sichtbarkeit von Forschungsprojekten ($F(5, 714) = 7.19$, $p < .001$) ist bei allen Fächern ein oft genutztes Bewertungskriterium. Besonders hoch fällt der Anteil an Nennungen in der Politikwissenschaft (77 Prozent) aus, während die Werte mit 41 Prozent in den Wirtschaftswissenschaften und 48 Prozent im Fach Sozialwesen signifikant niedriger ausfallen. Der Anteil bei den übrigen Fächern bewegt sich um die 60 Prozent. Abschließend zeigt sich, dass eigene Erfahrungen ($F(5, 714) = 2.93$, $p = .012$) ins-

besondere in der Politikwissenschaft (44 Prozent) einen relevanten Bewertungsmaßstab für die Forschungsleistung anderer Standorte darstellen. Signifikant niedriger fallen die Werte in der Erziehungswissenschaft (20 Prozent) und im Fach Sozialwesen (22 Prozent) aus. Die Werte der anderen Fächer liegen zwischen 26 und 34 Prozent. Auch hier könnte die Fachgröße als Erklärung eine Rolle spielen: Die geringere Größe der Fachcommunity bzw. Anzahl politikwissenschaftlicher Studienstandorte ermöglicht eher einen direkten Austausch (z.B. in Form von Forschungsk Kooperationen). Die gesammelten Erfahrungswerte können dann zur Beurteilung der verschiedenen Standorte herangezogen werden.

Die bis hierhin dargestellten Ergebnisse beziehen sich auf die Bewertungsmaßstäbe, welche die Professorinnen und Professoren zur Beurteilung der Lehre und Forschung von Hochschulstandorten sowie von einzelnen Studiengängen heranziehen. Diese sagen allerdings noch nichts darüber aus, ob die herangezogenen Informationen als ausreichend oder geeignet angesehen werden. Möglich ist durchaus, dass die genannten Kriterien nur deswegen herangezogen werden, weil sie die einzigen verfügbaren Informationsquellen darstellen und andere, eventuell geeignetere Bewertungskriterien nicht zur Verfügung stehen. Daher wurde in der Befragung – mit einem Fokus auf den von außen schwer einzuschätzenden Bereich der Lehre – auch erfragt, welche Kriterien sich die Professorinnen und Professoren zur Beurteilung von Studienstandorten ihres Faches wünschen. Die Anteilswerte der Nennungen für die Gesamtstichprobe sind in Abbildung 2 dargestellt. Es zeigt sich dabei, dass von über zwei Drittel der Befragten Informationen über die personelle Infrastruktur des Studienstandorts gewünscht werden. Zur Beurteilung herangezogen wurde die personelle Infrastruktur dagegen von nur etwa 36 Prozent der befragten Professorinnen und Professoren – d. h. ein großer Teil der Befragten hätte zwar gerne Informationen über die personelle Infrastruktur zur Beurteilung von Studienstandorten, aber bekannt, und damit als Bewertungsmaßstab verfügbar, waren sie den Befragten oftmals nicht.

Abbildung 2: Antworthäufigkeiten gewünschte Informationen zur Einschätzung eines Studienstandorts (n = 613)



Quelle: Eigene Berechnungen

Ähnlich oft wie die personelle Infrastruktur wurden mit einem Anteil von 62 Prozent Informationen über die Studieninhalte als Bewertungsmaßstab für andere Studienstandorte gewünscht. Diese dürften allerdings – sofern man nicht selbst an einem Standort tätig war – in den wenigsten Fällen in ausreichendem Maße vorhanden sein.

Interessanterweise werden von knapp 60 Prozent der Befragten auch Informationen über die Forschungsleistung als Beurteilungskriterium für den Studienstandort gewünscht. Gute Forschungsarbeit scheint also für einen großen Teil der Befragten auch ein Kriterium für einen guten Studienstandort zu sein. Gute Studienbedingungen werden aus dieser Perspektive folglich nicht losgelöst von guter Forschung gesehen.

Die Reputation von Kolleginnen und Kollegen ist nicht nur das Kriterium, das mit etwa 59 Prozent am häufigsten zur Beurteilung der Lehre eines Hochschulstandorts herangezogen wurde, sondern es werden von gut 52 Prozent der Befragten auch

Informationen hierzu zur Bewertung eines Studienstandorts gewünscht. Dennoch deutet die geringe Differenz von sieben Prozent zwischen den beiden Anteilswerten darauf hin, dass einige Befragte die Reputation von Kolleginnen und Kollegen nur als Bewertungskriterium herangezogen haben, weil andere, möglicherweise bessere Informationen nicht verfügbar waren. Andernfalls müsste der Anteil der gewünschten Informationen den Anteil der tatsächlich herangezogenen Informationen übersteigen.

Als weitere Kriterien zur Beurteilung eines Studienstandorts werden von je mehr als einem Drittel der befragten Professorinnen und Professoren Informationen über die materielle Infrastruktur (39 Prozent), internationale Kooperationen (38 Prozent), Arbeitsmarktchancen/Verbleibstudien (36 Prozent) sowie über das Theorie-Praxis-Verhältnis (35 Prozent) gewünscht. Auch Informationen über Abbrecherquoten und die internationale Ausrichtung werden mit etwa je 30 Prozent relativ oft genannt. Die Studiendauer (24 Prozent) sowie die Reputation des Standorts in der Wirtschaft (21 Prozent) werden nur von etwas weniger als einem Viertel der Befragten als Informationsquelle gewünscht. Zwar gibt nur knapp 16 Prozent der Befragten an, dass die an den Standort vergebene Lehrpreise eine gewünschte Information zur Einschätzung eines Studienstandorts darstellen; der Wert liegt damit aber deutlich höher als der Anteil an Befragten, die Lehrpreise tatsächlich als Bewertungsmaßstab für die Lehre eines Hochschulstandorts herangezogen haben (knapp 2 Prozent). Lehrpreise würden demnach öfter als Bewertungsmaßstab herangezogen werden, wenn diese in der jeweiligen Community sichtbarer wären.

Wie die Ergebnisse zeigen, entsprechen die zur Beurteilung (der Lehre und der Studiengänge) eines Hochschulstandorts herangezogenen Kriterien nicht immer den favorisierten, gewünschten Informationen. Diese sind in den entsprechenden Fällen scheinbar nicht ausreichend verfügbar oder für Außenstehende sichtbar, so dass zur Beurteilung anderer Standorte auf alternative Bewertungsmaßstäbe zurückgegriffen wird oder werden muss. Festzuhalten bleibt aber, dass sowohl bei den tatsächlich herangezogenen Bewertungsmaßstäben – für Forschung wie für Lehre – als auch bei den gewünschten Informationen personenbezogene Kriterien wie die personelle Infrastruktur und die Reputation von Kolleginnen und Kollegen eine entscheidende Rolle spielen. Dies scheint in besonderem Maße für Fächer mit einer kleineren fachspezifischen Community (hier Politikwissenschaft) zuzutreffen. Was der besondere Stellenwert personenbezogener Kriterien für die Beurteilung der Lehre und der Forschung von Hochschulstandorten und letztlich für die Reputation eines Hochschulstandorts bedeutet, wird nun in der nachfolgenden abschließenden Diskussion erörtert.

5 Fazit

Das wesentliche Kriterium, auf das sich Peers bei der Bewertung der Forschung und der Lehre eines sozialwissenschaftlichen Hochschulstandorts sowie bei der Bewertung von Studiengängen stützen, ist, neben der generellen personellen Infrastruktur, die Reputation von Kolleginnen und Kollegen. Die Einschätzung der Leistungsfähigkeit eines Fachbereichs in Forschung und Lehre sowie der Qualität eines Studiengangs ist damit in hohem Maße von der Reputation der dort tätigen Professorinnen und Professoren abhängig. Dabei kann davon ausgegangen werden, dass aufgrund der höheren Sichtbarkeit der Forschungsleistung gegenüber den nach außen weniger sichtbaren Leistungen im Bereich Lehre die durch die Peers wahrgenommene Reputation von Wissenschaftlerinnen und Wissenschaftlern primär auf deren Forschungsleistung gründet. Die Reputation einer Wissenschaftlerin oder eines Wissenschaftlers wäre damit überwiegend mit der Forschungsreputation gleichzusetzen – eine Annahme, die sich auch in den Gesprächen mit Fachvertretern im Rahmen der fachspezifischen Expertenworkshops des Projektes *QualRep* bestätigte. Eine hohe Forschungsreputation kann wiederum in Form eines Halo-Effektes auf andere Tätigkeitsfelder abstrahlen. So wird Forschenden, die sich beispielsweise durch herausragende Publikationen und Keynotes bei bedeutenden Tagungen eine hohe Forschungsreputation aufgebaut haben, aus Mangel an alternativen Informationen vielfach automatisch auch eine gute Lehrleistung attestiert werden.

Wenn die Reputation von Hochschulstandorten in hohem Maße von der Reputation der dort tätigen Wissenschaftlerinnen und Wissenschaftlern herrührt, gilt analog wie für den einzelnen Forscher, dass eine hohe Forschungsreputation eines Hochschulstandorts auf die Lehrreputation eines Standorts abstrahlen kann. Auch hier wird dann in diesem Sinne gute Forschung mit guter Lehre gleichgesetzt. Nun könnte man annehmen, dass ein Hochschulstandort entsprechend obiger Überlegungen Reputation verliert, wenn die hoch angesehenen Professorinnen und Professoren, die den Hochschulstandort maßgeblich prägen, diesen irgendwann verlassen. Dies ist in der Form nicht unbedingt der Fall. Vielmehr kommen hier wieder spezifische Besonderheiten des Konstrukts Reputation wie die nach- und abstrahlende oder auch kumulierende Wirkung von Reputation zum Tragen, die dafür sorgen, dass sich einmal aufgebaute Reputation im Zeitverlauf eher verfestigt als dass sie wieder verschwindet (vgl. *Fidler/Parsons 2008; Rauhvargers 2011*). Zudem sollte berücksichtigt werden, dass die Reputation von Professorinnen und Professoren zwar den am häufigsten genannten Bewertungsmaßstab für die Beurteilung von Hochschulstandorten darstellt, daneben aber auch andere Kriterien wie z.B. eigene Erfahrungen, Absolventengespräche, Publikationen, etc. eine Rolle spielen und für den Aufbau von Reputation eines Standorts maßgeblich sein können. Auch weitere, in der vorliegenden Studie nicht erhobene Merkmale eines Hochschulstandorts wie das Gründungsjahr (Tradition), die geo-

graphische Lage oder der generelle Hochschultyp (z.B. angebotene Studienfächer) können einen Einfluss auf die wahrgenommene Reputation eines Standorts haben (vgl. *Cyrenne/Grant 2009; Honecker 2011; Ridder 2010*). Nicht zuletzt gibt es Befunde dafür, dass Reputationsmessungen von Hochschulen stark durch die Stichprobenstruktur (Fachzugehörigkeit und Herkunft der Befragten) verzerrt sein können (*Federkeil 2009*).

Als Ergebnis der hier vorgestellten Analysen kann aber festgehalten werden, dass die Reputation eines Hochschulstandorts als Indikator für die Qualität von diesem eher ungeeignet erscheint, da die Reputation eines Standorts primär auf die Forschungsreputation der dort tätigen Professorinnen und Professoren zurückzuführen ist und andere wichtige qualitätsrelevante Aspekte damit in den Hintergrund gerückt werden. Dies gilt insbesondere für solche Aspekte, die keinen (direkten) Bezug zum Bereich Forschung aufweisen wie die Qualität der Lehre. Auch eine differenziertere Beurteilung, z.B. nach Standorten mit einer hohen Forschungsreputation auf der einen und Standorten mit einer hohen Lehrreputation auf der anderen Seite, ist nicht zielführend, da hier mit Überlagerungen der Lehrreputation durch die Forschungsreputation eines Standorts im Sinne eines Halo-Effektes zu rechnen ist. Der Nutzung des mehrdimensionalen Konstrukts der Reputation eines Hochschulstandorts als Qualitätsindikator wäre damit die Verwendung möglichst objektiver und eindimensionaler Indikatoren vorzuziehen (vgl. hierzu auch die Empfehlungen von *Kroth/Daniel 2008*).

In der Konsequenz wäre die Aussagekraft des in Rankings häufig verwendeten Reputationsindikators zumindest dann zu hinterfragen, wenn diese als Orientierung für Studierende bzw. Studieninteressierte dienen sollen. Gleichzeitig ist festzustellen, dass Rankings grundsätzlich primär auf Forschungsindikatoren zurückgreifen und kaum über systematisch-vergleichbare Informationen zur Qualität der Lehre an Standorten und in verschiedenen Fächern verfügen. Die Reputation eines Standorts oder Studiengangs bietet hier leider keine adäquate Alternative. Darüber hinaus sind die Ergebnisse aber auch für etablierte Qualitätssicherungssysteme im Hochschulbereich, allen voran Akkreditierungsverfahren, relevant, die in erster Linie auf Begutachtungen durch Peers aufbauen. Da eine Anonymisierung von Standorten und Studiengängen hier kaum möglich ist, wäre auch hier ein Bias in den Beurteilungen der Studiengänge auszugehen, der durch die (Forschungs-)Reputation der am Studiengang beteiligten Hochschullehrenden ausgelöst wird. Bislang liegen aber kaum Untersuchungen hierzu vor, die diese Zusammenhänge systematisch prüfen.

Einschränkungen in der Aussagekraft der vorliegenden Studie ergeben sich daraus, dass die einzelnen Dimensionen, aus denen sich die Konstrukte der Forschungs- und Lehrreputation zusammensetzen, hier sicherlich nicht erschöpfend erfasst wurden. Wie bereits angesprochen spielen bei der Beurteilung, ob ein Standort als heraus-

ragend gut in Forschung oder Lehre wahrgenommen wird, auch eine Vielzahl anderer Aspekte wie z.B. die Tradition eines Standorts eine Rolle. Eine umfassende Analyse, welche Dimensionen mit welchem Gewicht in die Reputation eines Standorts sowie in die Reputation von Wissenschaftlerinnen und Wissenschaftlern einfließen, wäre daher wünschenswert.

Aus forschungspraktischen Gründen musste sich die Erhebung der gewünschten Informationen zur Beurteilung eines Standorts auf den Lehrbereich beschränken (Studienstandort). Interessant wäre hier sicherlich auch, den Fokus auf den Bereich Forschung zu legen und zu ermitteln, welche Informationen aus Sicht von Professorinnen und Professoren zur Bewertung der Forschungsleistung eines Standorts notwendig bzw. wünschenswert sind. Schließlich muss festgehalten werden, dass die hier getätigten Schlussfolgerungen auf Basis von überwiegend rein deskriptiven Analysen erfolgten und damit eher den Charakter von ersten datengestützten Annahmen haben, die statistisch aber nicht abgesichert sind.

Zukünftig wäre weiter nach belastbaren Indikatoren für die Bewertung von Studiengängen zu fragen, nachdem sich Reputation hier als wenig hilfreich erwiesen hat. Die Leistungsfähigkeit von Akkreditierungen und anderen Verfahren der Qualitätssicherung ist bislang noch zu wenig untersucht. Hier ist die Hochschulforschung gefordert, wissenschaftlichen Kriterien genügend Kriterien zu entwickeln und Wege einer vergleichenden Qualitätseinschätzung von Studiengängen aufzuzeigen. Die hierfür bislang z.B. in Rankings genutzten Indikatoren werden der Komplexität des Gegenstands nicht gerecht.

Literatur

Armbruster, Bernd; Lederbogen, Utz (2010): Imagestudien als Evaluierungsinstrument strategischer Hochschulkommunikation. In: Vom Image zur Reputation Publikationen zur Hochschul-PR, Bd. 4, S. 35–54 Retrieved from http://www.bundesverband-hochschulkommunikation.de/fileadmin/user_upload/IQ/Projekte/Kleine_Datei_Image_und_Reputationsmanagement.pdf

Cyrenne, Philippe; Grant, Hugh (2009): University decision making and prestige: an empirical study. In: Economics of education review, 28 (2), S. 237–248

Enders, Jürgen (1998): Berufsbild der Hochschullehrer. In: Teichler, Ulrich u. a. (Hrsg.): Brennpunkt Hochschule. Neuere Analysen zu Hochschule, Beruf und Gesellschaft. Frankfurt am Main u. a., S. 55–78

Federkeil, Gero (2009): Reputation indicators in rankings of higher education institutions. In: Kehm, Barbara; Stens, Björn (Hrsg.): University Rankings, Diversity, and the New Landscape of Higher Education. Rotterdam/Taipeh, S. 19–34

Federkeil, Gero (2013): Internationale Hochschulrankings. Eine kritische Bestandsaufnahme. Beiträge zur Hochschulforschung, 35 (2), S. 34–48

Fidler, Brian; Parsons, Christine (2008): World University Ranking Methodologies: Stability and Variability. In: Higher Education Review, 40 (3), S. 15–34

Honecker, Patrick (2011): Image- und Reputationsmanagement als strategischer Prozess. Überlegungen zur Organisation von Hochschulkommunikation. In: Wissenschaftsmanagement 17 (5), S. 17–22

Kroth, Anna; Daniel, Hans-Dieter (2008): Internationale Hochschulrankings. Ein methodenkritischer Vergleich. In: Zeitschrift für Erziehungswissenschaft 11 (4), S. 542–558

Kühl, Stefan (2015): Reputation. Zur Funktion des Strebens nach Anerkennung in der Wissenschaft. In: Forschung & Lehre 22 (10), S. 804–806

Minssen, Heiner; Molsich, Beate; Wilkesmann, Uwe; Andersen, Uwe (2003): Kontextsteuerung von Hochschulen? Folgen der indikatorisierten Mittelzuweisung. Abhandlungen zu Bildungsforschung und Bildungsrecht, Bd. 12. Berlin

Rauhvargers, Andrejs (2011): Global university rankings and their impact. European University Association (EUA). Retrieved from http://www.eua.be/pubs/Global_University_Rankings_and_their_Impact.pdf

Ridder, Daniela de (2010): Reputationsmanagement an Hochschulen. Den guten Ruf strategisch planen. In: Wissenschaftsmanagement 16 (4), S. 53–54

Schmidt, Bernhard; Weishaupt, Horst (2008): Forschung und wissenschaftlicher Nachwuchs. In: Tillmann, Klaus-Jürgen (Hrsg.): Datenreport Erziehungswissenschaft 2008, S. 113–138

Schmidt-Hertha, Bernhard; Thalhammer, Veronika; Müller, Margaretha (2015): Studiengangs-Reputation als Indikator für Lehrqualität? In: Langemeyer, Ines u.a. (Eds.): Epistemic and Learning Cultures at the University of the 21st Century. Weinheim, S. 72–84

Statistisches Bundesamt (2015): Personal an Hochschulen 2014. Fachserie 11, Reihe 4.4. Wiesbaden: Statistisches Bundesamt

Statistisches Bundesamt (2016): Finanzen der Hochschulen 2014. Fachserie 11, Reihe 4.5. Wiesbaden: Statistisches Bundesamt

Statistisches Bundesamt (2017): Personal an Hochschulen 2016. Fachserie 11, Reihe 4.4. Wiesbaden: Statistisches Bundesamt

Trute, Hans-Heinrich, Pilniok, Arne (2009): Von der Ordinarien- über die Gremien- zur Managementuniversität? Veränderte Governance-Strukturen der universitären Forschung und ihre normativen Konsequenzen. In: Jansen, Dorothea u.a. (Hrsg.): Neue Governance für die Forschung. Tagungsband anlässlich der wissenschaftspolitischen Tagung der Forschergruppe „Governance der Forschung“, Berlin, 14.–15. März 2007

(S. 21–35). 1. Aufl. Baden-Baden: Nomos. (Interdisziplinäre Schriften zur Wissenschaftsforschung, Bd. 5)

Wissenschaftsrat (Hrsg.) (2008): Empfehlungen zur Qualitätsverbesserung von Lehre und Studium. Retrieved from <http://www.wissenschaftsrat.de/download/archiv/8639-08.pdf>

Wolbring, Tobias (2013): Fallstricke der Lehrevaluation. Möglichkeiten und Grenzen der Messbarkeit von Lehrqualität. Frankfurt am Main/New York

Manuskript eingereicht: 30.09.2017
Manuskript angenommen: 21.02.2018

Anschriften der Autoren:

Dr. Lorenz Tafertshofer
Wissenschaftlicher Mitarbeiter
Universität Tübingen
Institut für Erziehungswissenschaft
Münzgasse 11
72070 Tübingen
E-Mail: lorenz.tafertshofer@gmx.de

Evamaria Werner M.A.
Wissenschaftliche Mitarbeiterin
Universität Tübingen
Institut für Erziehungswissenschaft
Münzgasse 11
72070 Tübingen
E-Mail: evamaria.werner@uni-tuebingen.de

Prof. Dr. Bernhard Schmidt-Hertha
Professor für Erziehungswissenschaft mit Schwerpunkt berufliche und betriebliche Weiterbildung
Universität Tübingen
Institut für Erziehungswissenschaft
Münzgasse 11
72070 Tübingen
E-Mail: Bernhard.Schmidt-Hertha@uni-tuebingen.de

Buchvorstellung

Otto Hüther, Georg Krücken (2018): Higher Education in Germany – Recent Developments in an International Perspective. Wiesbaden: Springer, ISBN: 978-3-319-61478-6, 263 Seiten

Angeregt durch das wachsende internationale Interesse am deutschen Hochschulsystem und den Veränderungen, denen es in den letzten Dekaden unterworfen ist, legen Otto Hüther und Georg Krücken mit dieser Monographie eine umfassende Darstellung der Entwicklungen der letzten Dekaden vor. Aus einer interdisziplinären, historisch und international eingebetteten sowie theoriegeleiteten Perspektive werden zunächst die Reformbemühungen der letzten Jahrzehnte geschildert und innerhalb allgemeinerer und transnationaler Prozesse des sozialen Wandels verortet. Die folgenden Kapitel konzentrieren sich auf Veränderungen auf verschiedenen Ebenen: auf der Makro-Ebene (u.a. Veränderungen in der quantitativen und strukturellen Konfiguration der Hochschullandschaft durch Bildungsexpansion und steigende Studierendenzahlen), der Meso-Ebene (vor allem den Governance-Strukturen und Organisationsformen) und der Mikro-Ebene (u.a. wissenschaftliche Karriereverläufe, Hochschulprofessionen und soziale Chancengleichheit im höheren Bildungswesen). In einem abschließenden Kapitel werden die Erkenntnisse zusammengefasst und reflektiert, inwieweit es sich bei den Wandlungsprozessen um radikale Brüche oder eher um inkrementelle Veränderungen handelt.

Hinweise für Autorinnen und Autoren

Konzept:

Die Zeitschrift „Beiträge zur Hochschulforschung“ bietet Hochschulforschenden 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, Überblicksartikeln und Einblicken in die Praxis 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 externe Begutachtung durch anonyme Sachverständige (double blind). 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 Autorinnen und Autoren werden über das Ergebnis schriftlich informiert und erhalten gegebenenfalls Hinweise zur Überarbeitung.

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 Graphiker erstellt.

Bitte beachten Sie in jedem Fall bei Einreichung eines Manuskripts die ausführlichen verbindlichen Hinweise für Autoren unter <http://www.bzh.bayern.de>.

Kontakt:

Dr. Lydia Hartwig

E-Mail: Beitraege@ihf.bayern.de

Aus dem Inhalt

Michelle Mallwitz, Philipp Adler, Guido Benzler: Der Einfluss von Evaluierungen auf Forschungseinrichtungen am Beispiel der Leibniz-Gemeinschaft

Bernd Kleimann, Malte Hückstädt: Auswahlkriterien in Berufungsverfahren: Universitäten und Fachhochschulen im Vergleich

Heike Behle, Sunil Maher: Measuring "Teaching Excellence" and "Learning Gain" in the United Kingdom

Lorenz Tafertshofer, Evamaria Werner, Bernhard Schmidt-Hertha: Grundlagen der Reputation von Studienstandorten: Bewertungsmaßstäbe für die Qualität von Hochschulstandorten und Studiengängen aus der Sicht von sozialwissenschaftlichen Professorinnen und Professoren