

Drs. 2702-12
Köln 02 11 2012

Pilotstudie Forschungsrating Anglistik und Amerikanistik

Abschlussbericht der Bewertungsgruppe

INHALT

Vorbemerkung	5
Kurzfassung	6
Präambel	10
A. Erfahrungen und Empfehlungen	12
A.I Fachspezifische Operationalisierung	15
I.1 Definition des Fachs und der Teilbereiche	15
I.2 Definition der Kriterien und Festlegung der Daten	18
I.3 Bedingungen der Teilnahme	26
I.4 Vorbereitung der Erhebung	28
A.II Erhebung der Daten bei den Einrichtungen	30
II.1 Prinzipien der Datenerhebung	30
II.2 Organisation und Ablauf	32
II.3 Datenaufbereitung	37
A.III Bewertung	38
III.1 Definition der Bewertungsskala	38
III.2 Verfahren der Bewertung	38
III.3 Bewertungsmatrix	42
A.IV Veröffentlichung der Ergebnisse	42
B. Zusammenfassende Bewertung und Folgerungen	45
B.I Zusammenfassende Bewertung von Aufwand und Ertrag	45
B.II Weiteres Vorgehen	48
Abkürzungsverzeichnis	49
Literaturverzeichnis	50
Abbildungsverzeichnis	51

Vorbemerkung

Das Forschungsrating im Fach Anglistik und Amerikanistik geht auf die Empfehlungen des Wissenschaftsrats zum Forschungsrating vom Mai 2008 zurück. Darin hatte sich der Wissenschaftsrat auf Basis der Ergebnisse einer ersten Pilotstudie in den Fächern Chemie und Soziologie dafür ausgesprochen, das Verfahren anhand eines technik- und eines geisteswissenschaftlichen Fachs schrittweise weiterzuentwickeln und die Vorschläge der Steuerungsgruppe zur Optimierung des Verfahrens zu erproben. Nach Abschluss der Pilotstudie zur Weiterentwicklung des Forschungsratings wird der Wissenschaftsrat Empfehlungen dazu abgeben, ob und gegebenenfalls in welcher Form und durch wen ein Forschungsrating in Deutschland regelmäßig und in der Breite der Fächer durchgeführt werden soll. In diese Empfehlungen werden auch die Erfahrungen aus dem Forschungsrating Anglistik und Amerikanistik eingehen.

Im vorliegenden Abschlussbericht fasst die Bewertungsgruppe Anglistik und Amerikanistik ihre Erfahrungen aus der Pilotstudie zusammen und gibt Empfehlungen dazu ab, wie das Verfahren weiterentwickelt werden sollte, falls der Beschluss gefällt wird, es nach einigen Jahren erneut, gegebenenfalls auch in anderen Geisteswissenschaften durchzuführen.

Der vorliegende Abschlussbericht wurde von der Bewertungsgruppe Anglistik und Amerikanistik am 02. November 2012 verabschiedet.

Kurzfassung

Der Wissenschaftsrat hat im Mai 2008 beschlossen, das von ihm in den Fächern Chemie und Soziologie entwickelte Forschungsrating anhand eines technikwissenschaftlichen und eines geisteswissenschaftlichen Faches weiterzuentwickeln. |¹ In den Geisteswissenschaften wurde das Verfahren anhand der Anglistik und Amerikanistik erprobt. Damit liegt nun die vierte Pilotstudie zur Weiterentwicklung des Forschungsratings vor. Im Januar 2013 wird der Wissenschaftsrat Empfehlungen abgeben, ob und gegebenenfalls in welcher Form und durch wen ein Forschungsrating in Deutschland regelmäßig und in der Breite der Fächer durchgeführt werden soll. In diese Empfehlungen werden auch die Erfahrungen aus dem Forschungsrating Anglistik und Amerikanistik eingehen. Die Pilotstudie Forschungsrating im Fach Anglistik und Amerikanistik schließt methodisch an die im Rahmen der Pilotstudien in Chemie, Soziologie sowie Elektrotechnik und Informationstechnik entwickelten Grundzüge des Verfahrens an.

Die Durchführung des Forschungsratings in einem Fach lässt sich in vier Phasen unterteilen: 1. Fachspezifische Operationalisierung, 2. Erhebung der Daten bei den Einrichtungen, 3. Bewertung der Daten durch die Bewertungsgruppe, 4. Ergebnisveröffentlichung und Empfehlungen zum Verfahren. Der vorliegende Abschlussbericht greift jeden dieser Schritte auf und erläutert das jeweils konkrete Vorgehen. Auf der Basis der Erfahrungen des Forschungsratings Anglistik und Amerikanistik werden darüber hinaus Empfehlungen für künftige Forschungsratings formuliert.

1 – Fachspezifische Operationalisierung:

– Die von der Bewertungsgruppe vorgenommene Definition des Fachs Anglistik und Amerikanistik und seiner Teilbereiche hat sich grundsätzlich bewährt und als handhabbar erwiesen. Um Veränderungen aufzeigen zu können, soll-

¹ Die „Empfehlungen zum Forschungsrating“ (Drs. 8485-08) sind im Internet veröffentlicht unter <http://www.wissenschaftsrat.de/download/archiv/8485-08.pdf>

ten bei einer erneuten Bewertung der Anglistik und Amerikanistik die Teilbereichsdefinitionen (Englische Sprachwissenschaft, Anglistik: Literatur- und Kulturwissenschaft, Amerikastudien, Fachdidaktik Englisch) wieder verwendet werden. Damit Fehler bei der Selbstzuordnung der Wissenschaftlerinnen und Wissenschaftler zu den definierten Teilbereichen vermieden werden, sollten diese durch die Gutachterinnen und Gutachter fachlich geprüft und den Einrichtungen ggf. Korrekturen vorgeschlagen werden. Eine Herausforderung bleibt zudem vorerst eine adäquate Berücksichtigung von interdisziplinär Forschenden.

5
10
15
_ Um die unterschiedlichen Aufgaben und Profile der Einrichtungen widerzuspiegeln sowie die Stärken und Schwächen der Einrichtungen aufzuzeigen, wurden die Forschungsleistungen auch im Forschungsrating Anglistik und Amerikanistik nach mehreren Kriterien (Forschungsqualität, Reputation, Forschungsermöglichung und Transfer an außeruniversitäre Adressaten) bewertet, denen jeweils zwei bis drei Bewertungsaspekte zugeordnet waren. Diese wurden größtenteils durch qualitative Informationen operationalisiert. Die erhobenen Rahmeninformationen zu personellen Ressourcen und Lehrauslastungen erlaubten im Rahmen des Bewertungsprozesses eine Kontextualisierung der Daten zu Forschungsaktivitäten.

20 2 – Erhebung der Daten bei den Einrichtungen:

Die Publikations- und die Datenerhebung erfolgten nach dem *current-potential*-Prinzip (Leistung der zum Stichtag 31.12.2012 an der betreffenden Einrichtung beschäftigten Personen für den zurückliegenden 7-Jahres-Zeitraum). Da jedoch nicht alle relevanten Daten an den Einrichtungen stichtagbezogen vorlagen, wurde ein Teil der Informationen nach dem *work-done-at*-Prinzip erhoben. (Leistung aller an der betreffenden Institution im 7-Jahres-Zeitraum vom 01.01.2004 bis 31.12.2010 beschäftigten Wissenschaftlerinnen und Wissenschaftler). Somit lag bei der Datenerhebung ein so genannter „hybrider“ Ansatz von *current-potential* und *work-done-at* zugrunde.

25
30 Die Datenerhebung erfolgte in drei Schritten: 1. Personenerhebung, 2. Publikationserhebung und 3. Hauptdatenerhebung. Im ersten Schritt wurden Wissenschaftlerinnen und Wissenschaftler, die in der Anglistik und Amerikanistik forschend tätig sind, differenziert nach Personalkategorien erfasst und von den Einrichtungen den vier definierten Teilbereichen zugeordnet. Anschließend wurden die Einrichtungen gebeten, pro Professur drei exemplarische Publikationen aus dem Erhebungszeitraum einzureichen. Im Zuge der sich daran anschließenden Hauptdatenerhebung wurden alle übrigen bewertungsrelevanten Daten erhoben.

35
40 Mit Ausnahme der exemplarischen Publikationen, wurden die Daten bei den Einrichtungen mit Hilfe von Online-Fragebögen erhoben, die bereits im For-

schungsrating Elektrotechnik und Informationstechnik erfolgreich erprobt und in der Anglistik und Amerikanistik weiterentwickelt wurden.

3 – Bewertung der Daten durch die Bewertungsgruppe

Grundsätzlich haben sich das Verfahren und die gewählte Vorgehensweise der Bewertung in Form eines *informed-peer-review*-Verfahrens, wie bereits in den vorangegangenen Pilotstudien, bewährt. Die Bewertung wurde in drei Stufen vorgenommen: Zunächst ermittelten die jeweils zuständigen Gutachterinnen und Gutachter unabhängig voneinander ihre vorläufigen Bewertungsvoten. Anschließend wurden die vorläufigen Ergebnisse in teilbereichsspezifischen Gruppen beraten, wobei immer zwei Teilbereiche zusammen tagten (Anglistik: Literatur und Kulturwissenschaft zusammen mit Amerikastudien sowie Englische Sprachwissenschaft zusammen mit Fachdidaktik Englisch). In einem letzten Schritt wurden alle Bewertungen im Plenum zur Abstimmung gestellt.

Um die Verfasstheit des Faches angemessen zu berücksichtigen, wurden alle Kriterien auf der Ebene der Teilbereiche bewertet. Nach der ersten Sichtung der Daten und in Vorbereitung auf die Bewertungsphase fanden in Begleitung der Mitarbeiterinnen und Mitarbeiter der Geschäftsstelle des Wissenschaftsrats Treffen der Gutachterinnen und Gutachter der jeweiligen Teilbereiche statt, die der Entwicklung von teilbereichsspezifischen Bewertungsmaßstäben dienen. Dadurch wurde eine frühzeitige Auseinandersetzung mit dem Datenmaterial sichergestellt und ein für die einzelnen Teilbereiche angemessener Zugang zur Bewertung geschaffen. Diese Vorgehensweise hat sich bewährt und sollte auch künftig beibehalten werden, wobei großer Wert darauf zu legen ist, die Ergebnisse im Anschluss an diese teilbereichsspezifischen Treffen zusammenzuführen und die Bewertungsmaßstäbe in der Bewertungsgruppe gemeinsam zu definieren.

Die zur Bewertung vorliegenden Daten erwiesen sich als unterschiedlich aussagekräftig. Die für die Bewertung der Kriterien Forschungsqualität und Forschungsermöglichung erhobenen Daten lieferten eine solide und belastbare Grundlage. Die Bewertung der Kriterien Reputation und Transfer an außeruniversitären Adressaten war durch teilweise unvollständige Daten hingegen erschwert. Das Bewertungsmodell hat sich gleichwohl bewährt und sollte unter Berücksichtigung der empfohlenen Anpassungen beibehalten werden. Effizienzmaße wurden nicht berechnet. Zugleich waren die erhobenen Rahmeninformationen für die Einordnung und Kontextualisierung der übrigen Daten hilfreich. Die hohe Übereinstimmung der Bewertungen zwischen den Gutachterinnen und Gutachtern können als Beleg für die Verlässlichkeit des *informed-peer-review*-Verfahrens gewertet werden.

Wie auch in den vorangegangenen Pilotstudien besteht die Ergebnisveröffentlichung aus zwei Teilen, dem Ergebnisbericht und den einrichtungsbezogenen Ergebnisdarstellungen. Die Ergebnisse sind auch im Internet unter www.forschungsrating.de veröffentlicht und erlauben für die vier definierten Teilbereiche einen direkten Vergleich der Einrichtungen auf der Ebene der einzelnen Kriterien.

Die erfolgreich durchgeführte Pilotstudie Forschungsrating Anglistik und Amerikanistik zeigt, dass eine dem Fach angemessene vergleichende Bewertung von Forschungsleistungen in den Geisteswissenschaften im Allgemeinen und in der Anglistik und Amerikanistik im Besonderen grundsätzlich möglich ist. Das Forschungsrating ist geeignet, im Rahmen von Forschungsbewertung, den Spezifika der geisteswissenschaftlichen Forschungspraxis Rechnung zu tragen. Dies schlägt sich beispielsweise in der Entwicklung und Operationalisierung des Bewertungsmodells und der Bestimmung des Erhebungszeitraums nieder. Die nach Teilbereichen und Kriterien differenzierte Darstellungsweise bietet adressatengerechte Informationen.

Die Bewertungsgruppe befürwortet eine turnusmäßige Wiederholung des Forschungsratings in der Anglistik und Amerikanistik.

Präambel

Bei der Suche nach einem für die Pilotstudien des Forschungsratings repräsentativen geisteswissenschaftlichen Fach fiel nach verschiedenen Gesprächen mit dem Vorstand der Fachverbände, dem Deutschen Anglistenverband und der Deutschen Gesellschaft für Amerikastudien, die Wahl auf die Anglistik und Amerikanistik. Das Selbstverständnis dieses Faches sowie der Fachvertreterinnen und Fachvertreter ist durch die Beschäftigung mit den sich global verbreiteten und rezipierten Sprachen, Literaturen und Kulturen der anglophonen Welt geprägt von vielfältigen Kontakten und Kooperationen mit Institutionen in englischsprachigen Ländern und somit offen für einen internationalen Forschungswettbewerb. Die Kenntnis der seit vielen Jahren in Großbritannien und Nordamerika regelmäßig durchgeführten Evaluationsverfahren, die einerseits die Bewertung der Lehre durch Studierende, andererseits aber auch die Erfassung der Forschungsaktivität betrifft, hat die in geisteswissenschaftlichen Fächern vorhandene Skepsis gegenüber solchen Formen der Leistungsbemessung relativiert. Angesichts der neuerdings auch in Deutschland vorgenommenen Ranking-Studien kommerzieller Unternehmen erschien das von der Steuerungsgruppe des Wissenschaftsrates entwickelte Forschungsrating durch *"informed-peer-reviewers"* eine willkommene Gelegenheit, von den Gegebenheiten des Faches aus Kriterien für die Bewertung der Forschung zu erstellen und diese durch einen ersten Praxistest zu überprüfen und zur Diskussion zu stellen. Aus den von den Fachverbänden, dem Wissenschaftsrat und der DFG vorgeschlagenen Kandidatinnen und Kandidaten wurde eine Bewertungsgruppe gebildet, die die Beteiligung von Wissenschaftlerinnen und Wissenschaftler nach fachlicher und regionaler Verteilung sowie verschiedener Karrierestufen berücksichtigte. Bei der ersten Sitzung dieser Bewertungsgruppe im Februar 2011 wurden nach der Definition von vier Teilbereichen des Faches in Anglistik: Literatur- und Kulturwissenschaft, Amerikastudien, Englische Sprachwissenschaft und Fachdidaktik Englisch Kriterien für das Forschungsrating entwickelt, die den Gegebenheiten dieses geisteswissenschaftlichen Faches entsprechen. Aufgrund der ersten Rückmeldungen durch die Einrichtungen wurden Gespräche mit der Deutschen Gesellschaft für Fremdsprachenforschung geführt und die Zahl der Gutachterinnen und Gutachter von 15 auf 19 angehoben, so dass schließlich jeweils fünf Gutachterinnen und Gutachter aus den Bereichen Anglistik, Ame-

rikanistik und Englische Sprachwissenschaft sowie vier Englischdidaktikerinnen und Englischdidaktiker beteiligt waren. Im Laufe des im Oktober 2012 abgeschlossenen Bewertungsprozesses konnten Korrekturen und Verbesserungsvorschläge in Gesprächen mit den Fachverbänden, den Fachkoordinatorinnen und Fachkoordinatoren sowie Kolleginnen und Kollegen bei Präsentationen auf Mitgliederversammlungen berücksichtigt werden. Dabei wurde den Einrichtungen auch die Möglichkeit der Reaktion auf die Bewertungsergebnisse vor der Veröffentlichung zugesichert. Die in großer Übereinstimmung erzielten Ergebnisse geben ein differenziertes Bild des Faches Anglistik und Amerikanistik in Deutschland und bilden eine gute Basis für die Weiterentwicklung des Forschungsratings. Ein solches Vorgehen setzt den Ranking-Verfahren eine differenzierte und wissenschaftsgeleitete Methodik entgegen.

A. Erfahrungen und Empfehlungen

Die Entwicklung des Forschungsratings geht zurück auf die „Empfehlungen zu Rankings im Wissenschaftssystem Teil 1: Forschung“, ¹² die der Wissenschaftsrat im November 2004 verabschiedet hat. Darin spricht er sich für eine vergleichende Bewertung der Forschungsleistungen aus mit dem Ziel, die wissenschaftlichen Einrichtungen bei strategischen Entscheidungen zu unterstützen und durch mehr Transparenz den Wettbewerb zu fördern. Bei dem Forschungsrating handelt es sich um ein wissenschaftsgeleitetes Verfahren, das sich durch *informed-peer-review*, Mehrdimensionalität, fachspezifische Operationalisierung sowie, sofern vorhanden, die explizite Einbeziehung außeruniversitärer Forschungseinrichtungen auszeichnet.

Zur Erprobung des vorgeschlagenen Verfahrens beschloss der Wissenschaftsrat im Juli 2005, eine Pilotstudie Forschungsrating zunächst in den Fächern Chemie und Soziologie durchzuführen. Auf Grundlage der im Mai 2008 abgeschlossenen ersten Pilotstudie ¹³ in diesen beiden Fächern empfahl er, das Forschungsrating anhand eines technik- und geisteswissenschaftlichen Fachs sukzessive weiterzuentwickeln. Vor allem sollte in der Folgestudie geprüft werden, ob mittelfristig eine alle geeigneten Fächer erfassende, turnusmäßige Durchführung möglich ist.

Die Gesamtverantwortung für das Forschungsrating obliegt einer vom Wissenschaftsrat eingesetzten Steuerungsgruppe, in der neben Mitgliedern der Wissenschaftlichen Kommission des Wissenschaftsrats und weiteren Sachverständigen die großen Wissenschaftsorganisationen Hochschulrektorenkonferenz

¹² Die „Empfehlungen zu Rankings im Wissenschaftssystem Teil 1: Forschung“ (Drs. 6285-04) sind im Internet veröffentlicht unter: <http://www.wissenschaftsrat.de/download/archiv/6285-04.pdf>

¹³ Wissenschaftsrat: Pilotstudie Forschungsrating. Empfehlungen und Dokumentation, Köln 2008. Die „Empfehlungen zum Forschungsrating“ (Drs. 8485-08) sind auch im Internet verfügbar unter: <http://www.wissenschaftsrat.de/download/archiv/8485-08.pdf>

(HRK), Deutsche Forschungsgemeinschaft (DFG), Fraunhofer-Gesellschaft (FhG), Helmholtz-Gemeinschaft Deutscher Forschungszentren (HGF), Max-Planck-Gesellschaft (MPG) und Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL) durch ex officio-Mitglieder vertreten sind. Sechs Länder, der Bund und die Geschäftsstellen der großen Wissenschaftsorganisationen entsenden zudem Gäste in die Steuerungsgruppe. |⁴

Die Auswahl der Fächer erfolgte durch Beschluss des Wissenschaftsrates und war bestimmt durch das Anliegen, möglichst unterschiedliche Fächergruppen in die Pilotstudie einzubeziehen. Die definitive Festlegung der Fächer traf die Steuerungsgruppe nach folgenden Kriterien:

1 – Repräsentativität und Bedeutung: Das Fach sollte möglichst repräsentativ und bedeutend für den jeweiligen Wissenschaftsbereich sein.

2 – Größe: Das Fach sollte hinreichend groß sein, um ein angemessenes Verhältnis von Aufwand und Ertrag zu erzielen. Jedoch sollte das Fach auch nicht zu umfassend definiert werden, um eine ausreichende Homogenität der Qualitätsmaßstäbe zu gewährleisten und die Gutachterbelastung in Grenzen zu halten.

3 – Abgrenzbarkeit und Komplexität: Das Fach sollte in konsensfähiger Weise abgrenzbar sein, um eine Zuordnung von Forschungseinheiten ohne taxonomische Grundsatzdiskussionen in den jeweiligen Fachgemeinschaften zu ermöglichen.

4 – Interdisziplinarität: Es sollten Bezüge zu anderen Fächern bestehen, um den Umgang mit interdisziplinären Ansätzen im Rahmen des Forschungsratings weiterzuentwickeln.

5 – Außeruniversitäre Forschungseinrichtungen: In dem Fach sollten möglichst auch außeruniversitäre Forschungseinrichtungen bestehen, da die Weiterführung des Forschungsratings klären soll, wie diese besondere Organisationsform angemessen berücksichtigt werden kann. Zudem gehört die Einbeziehung der außeruniversitären Forschung zu den Vorzügen des Forschungsratings gegenüber den bestehenden nationalen und internationalen Rankings. Spezifische Regelungen sind dabei für Forschungseinheiten erforderlich, deren wissenschaftliche Leiterinnen und Leiter durch gemeinsame Berufung an einer außeruniversitären Einrichtung und einer Universität zugleich tätig sind.

|⁴ Die Mitgliederliste der Steuerungsgruppe ist im Internet veröffentlicht unter:

<http://www.wissenschaftsrat.de/arbeitsbereiche-arbeitsprogramm/forschungsrating/organisation/>

6 – Unterstützung durch Fachverbände: Es sollte möglichst Ansprechpartnerinnen und -partner in den fachnahen Verbänden geben, die das Verfahren begleiten und unterstützen.

5 Unter Abwägung dieser Kriterien hat sich die Steuerungsgruppe für die Anglistik und Amerikanistik als Pilotstudie im Bereich der Geisteswissenschaften entschieden.

10 Die Operationalisierung des Verfahrens und die Bewertung der Forschungsleistungen erfolgen jeweils fachspezifisch. Die Steuerungsgruppe setzt dazu eine Bewertungsgruppe mit Fachgutachterinnen und Fachgutachtern ein. Um die Verfahrenskohärenz zu gewährleisten, gehört jeder Bewertungsgruppe neben den fachbezogenen Gutachterinnen und Gutachtern jeweils ein fachfernes Mitglied der Steuerungsgruppe mit Gaststatus an.

15 Die Pilotstudie Forschungsrating im Fach Anglistik und Amerikanistik schließt methodisch an die Grundzüge des Verfahrens an, die im Rahmen des Forschungsratings Chemie, Soziologie und Elektrotechnik und Informationstechnik entwickelt wurden. Darüber hinaus greift sie die Empfehlungen der Steuerungsgruppe zur vergleichenden Forschungsbewertung in den Geisteswissenschaften |⁵ auf und bietet die Möglichkeit, die Empfehlungen zur Weiterentwicklung und Optimierung des Verfahrens umzusetzen und zu erproben.

20 Die Durchführung des Forschungsratings in einem Fach lässt sich in vier Phasen unterteilen:

1 – Fachspezifische Operationalisierung

2 – Erhebung der Daten bei den Einrichtungen

25 3 – Bewertung der Daten durch die Bewertungsgruppe

4 – Ergebnisveröffentlichung und Empfehlungen zum Verfahren

30 Für jede dieser Phasen werden im Folgenden Vorgehen und Erfahrungen des Forschungsratings Anglistik und Amerikanistik bilanziert und auf dieser Grundlage Empfehlungen für die Weiterentwicklung des Forschungsratings abgeleitet.

|⁵ Wissenschaftsrat: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, Köln 2010.

Die fachspezifische Anpassung des Verfahrens an die Anglistik und Amerikanistik umfasste die Definition des Fachs und der Teilbereiche, die Definition der Kriterien und die Festlegung der Daten, die Formulierung der Teilnahmebedingungen sowie die Vorbereitung der Erhebung.

1.1 Definition des Fachs und der Teilbereiche

Die Definition des Fachs und der Teilbereiche stellte den ersten Schritt der fachspezifischen Anpassung des Verfahrens dar und erfüllte drei Funktionen. Zum einen sollten die Einrichtungen eine Handreichung für die Identifikation der relevanten Teilbereiche und somit eine Entscheidungsgrundlage für die Teilnahme am Forschungsrating und für die Meldung der teilnehmenden Teilbereiche bekommen. Zum anderen sollte die Definition der Teilbereiche es erlauben, im Rahmen der Bewertungen ein differenziertes Profil, das Aussagen über Qualitätsunterschiede innerhalb einer Einrichtung ermöglicht, zu zeichnen. Darüber hinaus diente die Abgrenzung des Faches sowie die Definition der Teilbereiche der Identifikation der wichtigsten Kompetenzfelder, die in der Bewertungsgruppe vertreten sein sollten, um die Teilbereiche bewerten zu können.

Aus dem Vergleich der Fachsystematiken der Deutschen Forschungsgemeinschaft, des Statistischen Bundesamtes (Lehr- und Forschungsbereiche, Fachgebiete) sowie der Fachtraditionen und der Fakultätsorganisation ausgewählter Hochschulen leitete die Bewertungsgruppe vier Teilbereiche ab:

- _ Englische Sprachwissenschaft,
- _ Anglistik: Literatur- und Kulturwissenschaft,
- 25 _ Amerikastudien,
- _ Fachdidaktik Englisch.

Da das Forschungsrating vornehmlich eine institutionelle Bewertung im Blick hat und damit sich der Erhebungsaufwand in Grenzen hält, sollte eine unnötige Kleinteilung vermieden werden. Zudem würde eine personenbezogene Bewertung einer Veröffentlichung der Ergebnisse für die betreffenden Untereinheiten entgegenstehen. Um jedoch Leistungsunterschiede innerhalb einer Einrichtung sichtbar zu machen, sollten die Teilbereiche auch nicht zu groß sein. Darüber hinaus sollten die Teilbereiche so definiert werden, dass die Struktur der verschiedenen Standorte vergleichbar blieb. Zugleich sollten taktische Kalküle einzelner Einrichtungen, den jeweiligen Zuschnitt der Forschungseinheiten zu beeinflussen, verhindert werden.

Um eine Kleinteiligkeit zu vermeiden, wurde auf eine Abgrenzung weiterer Teilbereiche wie der „*Postcolonial Studies*“ sowie der „Mediävistik“ verzichtet, die nur an wenigen Einrichtungen vertreten sind. Um den inhaltlichen und fachpolitischen Entwicklungen im Fachgebiet der Anglistik und Amerikanistik angemessen Rechnung zu tragen, auch mit Blick auf die Bewertung der Publikationsleistungen, wurden „Anglistik: Literatur- und Kulturwissenschaft“ sowie „Amerikastudien“ als jeweils eigenständige Teilbereiche und in Abgrenzung zur „Englischen Sprachwissenschaft“ definiert.

Mit der oben beschriebenen Definition der Teilbereiche sollten die theoretischen, methodischen, kulturellen und regionalen Aspekte der Forschung in der Anglistik und Amerikanistik sowie der Zuschnitt der Fächer innerhalb der Einrichtungen berücksichtigt werden. Zugleich ermöglichte diese offene Definition den interdisziplinär forschenden Wissenschaftlerinnen und Wissenschaftlern sowie in der Allgemeinen und Vergleichenden Sprach-, Literatur- und Kulturwissenschaft Forschenden die Teilnahme am Forschungsrating. Grundsätzlich sollte das Verfahren auch künftig gegenüber Teilbereichen mit einer interdisziplinären Ausrichtung offen sein und diese angemessen abbilden.

Im Rahmen der ersten Pilotstudien Forschungsrating Chemie und Soziologie waren die Teilbereiche nicht vorgegeben worden. Vielmehr wurden die Forschungseinheiten von den einzelnen teilnehmenden Einrichtungen selbst festgelegt. Als Folge variierte die Größe der gemeldeten Forschungseinheiten erheblich und erschwerte so den Bewertungsprozess. Mit der Vereinheitlichung und Definition der Teilbereiche im Rahmen des Forschungsratings Anglistik und Amerikanistik kommt die Bewertungsgruppe den entsprechenden Empfehlungen der Steuerungsgruppe nach und trägt somit zu Weiterentwicklung und wesentlichen Verbesserung dieses Aspekts gegenüber der ersten Pilotstudie bei. Nicht zuletzt ermöglichte die klare Definition und Abgrenzung der Teilbereiche innerhalb eines Fachs eine begründete Zuordnung von Berichterstatterinnen und Berichterstattern. Zudem wurden dadurch Teilbereiche identifiziert, für die weitere Bewertungsgruppenmitglieder hinzugewonnen werden mussten. So erforderte beispielsweise die Definition der Fachdidaktik Englisch als eigenständiger Teilbereich eine entsprechende Erweiterung und Nachbesetzung der Bewertungsgruppe.

Als problematisch erwies sich zugleich, dass in einzelnen Fällen die Selbstzuordnung durch die Einrichtungen nicht immer transparent war und aus der Sicht der Gutachterinnen und Gutachter teilweise zu fehlerhaften und/oder strategisch nachteiligen Zuordnungen führte. Im ungünstigsten Fall könnte dies zu einer schlechteren Bewertung oder gar Nicht-Bewertbarkeit des Teilbereichs führen. Betroffen waren hier insbesondere die Forschenden der Fachdidaktik Englisch, die den Teilbereichen „Englische Sprachwissenschaft“ oder der „Anglistik: Literatur- und Kulturwissenschaft“ zugeordnet wurden. Die Ent-

scheidung über die Zuordnung zu den einzelnen Teilbereichen war für die Pädagogischen Hochschulen besonders erschwert, da an diesen Wissenschaftlerinnen und Wissenschaftler der Englischen Sprachwissenschaft, der Anglistik: Literatur- und Kulturwissenschaft sowie der Amerikastudien in der Fachdidaktik Englisch bzw. interdisziplinär forschen und lehren.

Eine Mehrfachzuordnung von Personen zu verschiedenen Teilbereichen war im Rahmen des Forschungsratings grundsätzlich nicht zugelassen. Wissenschaftlerinnen und Wissenschaftler, deren Professur keine eindeutige Denomination hat oder die Forschungsleistungen auf mehreren Gebieten erbringen, wurden gebeten, sich demjenigen Teilbereich zuzuordnen, in dem sie selbst ihren Forschungsschwerpunkt sehen. Im Rahmen der Haupterhebung wurde die Möglichkeit eingeräumt, auf Schwierigkeiten bei der Zuordnung hinzuweisen. Die Bewertungsgruppe hat in diesen Fällen die gesamten Forschungsleistungen durch Querverweise und mit Hilfe von Gutachtern der anderen zuständigen Teilbereiche differenziert bewertet. Besonders betroffen waren auch hier die Pädagogischen Hochschulen, die ihre Wissenschaftlerinnen und Wissenschaftler zunächst sowohl der Fachdidaktik Englisch als auch einem der anderen drei Teilbereiche zugeordnet hatten.

Eine weitere Schwierigkeit für die Einrichtungen stellte die Zuordnung von Forschenden dar, die sich keinem der von der Bewertungsgruppe definierten Teilbereiche eindeutig zuordnen ließen. Im ungünstigsten Fall führte dies dazu, dass einschlägig forschungsaktive Personen nicht am Forschungsrating teilnahmen. Besonders zu nennen ist hier die Forschungsrichtung der Anglistischen Mediävistik. Die Bewertungsgruppe sprach sich gegen eine Definition der „Anglistischen Mediävistik“ als eigenständigen Teilbereich aus, da diese an nur noch wenigen Standorten in Deutschland vertreten ist und die Vergleichsgrundlage somit sehr klein gewesen wäre. Vielmehr wurden die Einrichtungen gebeten, Wissenschaftlerinnen und Wissenschaftler, die auf dem Gebiet der Mediävistik forschen, demjenigen Teilbereich zuzuordnen, der ihrer Forschung inhaltlich am nächsten steht.

Eine besondere Schwierigkeit der Zuordnung zu einem Teilbereich bestand an den Pädagogischen Hochschulen, da die an diesen tätigen Professorinnen und Professoren stets gleichzeitig eine Fachwissenschaft und die Fachdidaktik vertreten. Im Forschungsrating Anglistik und Amerikanistik hat sich letztlich das Gros der an Pädagogischen Hochschulen Forschenden dem Teilbereich „Fachdidaktik Englisch“ zugeordnet.

Schließlich sollten auch Doppelmeldungen von Wissenschaftlerinnen und Wissenschaftlern im Rahmen des Forschungsratings vermieden werden. Sofern eine Person zum Stichtag 31.12.2010 von zwei Einrichtungen gemeldet wurde, wurden beide Einrichtungen um weitere Informationen zum Beschäftigungsstatus der betreffenden Person gebeten. Diese ermöglichten eine abschließende Klä-

18 rung, für welche Einrichtung die Meldung der betreffenden Person am sinn-
 vollsten ist. Grundsätzlich wurde vorgeschlagen, eine Person derjenigen Ein-
 richtung zuzuordnen, an der sie zum Stichtag 31.12.2010 tatsächlich tätig war.
 In einigen wenigen Fällen einer Doppelaaffiliation wurde im Einvernehmen mit
 5 den Hochschulen eine Zuordnung zu einer Einrichtung vorgenommen.

Die von der Bewertungsgruppe für die erste Pilotstudie vorgenommene Defini-
 tion des Fachs Anglistik und Amerikanistik und seiner Teilbereiche hat sich
 grundsätzlich bewährt und als handhabbar erwiesen. Um Veränderungen auf-
 zeigen zu können, sollten bei einer erneuten Bewertung der Anglistik und Ame-
 10 rikanistik die Teilbereichsdefinitionen wieder verwendet werden. Verbesse-
 rungspotenzial wird in der Berücksichtigung der Mediävistik sowie
 interdisziplinär forschender Teilbereiche gesehen. Darüber hinaus sollte im
 Rahmen der Abfrage die Definition der einzelnen Teilbereiche präzisiert wer-
 den, um fehlerhafte Selbstzuordnungen zu vermeiden (z. B. durch eine Auflis-
 15 tung der jeweils zugehörigen Forschungsfelder). Die Selbstzuordnungen sollten
 zudem durch die Gutachterinnen und Gutachter fachlich geprüft werden, um
 Einrichtungen ggf. Korrekturen vorschlagen zu können.

1.2 Definition der Kriterien und Festlegung der Daten

Um die unterschiedlichen Aufgaben der Einrichtungen widerzuspiegeln, die un-
 20 terschiedlichen Profile zu zeichnen sowie um die Stärken und Schwächen der
 Einrichtungen aufzuzeigen, werden die Forschungsleistungen im Forschungsra-
 ting des Wissenschaftsrats nach mehreren Bewertungskriterien bewertet. Um
 die Vergleichbarkeit der einzelnen Bewertungen zu erhöhen, werden die ein-
 zelnen Kriterien nach mehreren Bewertungsaspekten bewertet. Zugleich wird
 25 anhand der Bewertungsaspekte entschieden, welche Daten erhoben werden sol-
 len.

Die Grundlage für die Definition der Kriterien im Forschungsrating Anglistik
 und Amerikanistik bilden die Empfehlungen des Wissenschaftsrats zu Rankings
 im Wissenschaftssystem, Teil 1: Forschung (2004) |⁶, Empfehlungen der Steue-
 30 rungsgruppe zum Forschungsrating |⁷ sowie zur vergleichenden Forschungs-
 bewertung in den Geisteswissenschaften (2010). |⁸

|⁶ Die „Empfehlungen zu Rankings im Wissenschaftssystem Teil 1: Forschung“ (Drs. 6285-04) sind im In-
 ternet veröffentlicht unter: <http://www.wissenschaftsrat.de/download/archiv/6285-04.pdf>.

|⁷ Die „Empfehlungen zum Forschungsrating“ (Drs. 8485-08) sind im Internet veröffentlicht unter:
<http://www.wissenschaftsrat.de/download/archiv/8485-08.pdf>.

|⁸ Vgl. Wissenschaftsrat: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissen-
 schaften, Köln 2010, S. 8.

Entsprechend den o. g. Empfehlungen des Wissenschaftsrats wurden die Forschungsleistungen in den Fächern Chemie und Soziologie in drei Leistungsdimensionen (Forschung, Nachwuchsförderung und Transfer) nach insgesamt sechs Kriterien bewertet. |⁹ Im Rahmen der fachspezifischen Anpassung hat die

5 Bewertungsgruppe Elektrotechnik und Forschungsrating die vom Wissenschaftsrat empfohlenen Leistungsdimensionen beibehalten, jedoch die Zahl der Kriterien auf fünf reduziert. Die Reduktion betraf die Dimension Wissenstransfer.

Die zur vergleichenden Forschungsbewertung verwendeten Bewertungskriterien müssen grundsätzlich dazu geeignet sein, alle Forschungseinrichtungen in einem Fach nach einheitlichen Maßstäben bewerten zu können. Zudem sollten die Bewertungskriterien und die dazu zu erhebenden Daten derart beschaffen sein, dass sie zur Differenzierung der Bewertungen der einzelnen Einrichtungen beitragen. Die Definition der Kriterien erfordert eine intensive Auseinandersetzung im Vorfeld der eigentlichen Bewertung, um den Prozess des *informed-peer-review* nachvollziehbar gestalten zu können.

10

15

Im Zuge der Anpassung der Bewertungskriterien an die Forschungspraxis der Anglistik und Amerikanistik und unter Berücksichtigung der Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften definierte die Bewertungsgruppe Anglistik und Amerikanistik vier Bewertungskriterien: „Forschungsqualität“, „Reputation“, „Forschungsermöglichung“ sowie „Transfer an außeruniversitäre Adressaten“ und ordnete diesen Bewertungsaspekte und die jeweils zu erhebenden qualitativen und quantitativen Informationen und Daten zu. Aus dieser Zuordnung resultierte die so genannte „Bewertungsmatrix“, |¹⁰ die die Grundlage der Bewertung bildete. Auf eine separate Bewertung des Kriteriums „Nachwuchsförderung“ wurde verzichtet, vielmehr wurden die dazugehörigen Leistungen im Rahmen des Kriteriums „Forschungsermöglichung“ bewertet. Um eine Einordnung der Daten zu ermöglichen, wurden ausgewählte Rahmeninformationen zur Lehrauslastung sowie zu vorhandenen Ressourcen erhoben. Gegenüber den bisherigen Pilotstudien wurde auf die Bewertung der Kriterien „Effizienz“ sowie „Impact/Effektivität“ verzichtet. Diese dienten zur Bewertung des absoluten Beitrags einer Einrichtung zur Entwicklung des Fachgebiets (Impact/Effektivität) sowie des auf die Größe und die eingesetzten Ressourcen relativierten Beitrags (Effizienz). Der Verzicht auf die

20

25

30

|⁹ Ursprünglich waren vom Wissenschaftsrat neun Bewertungskriterien vorgeschlagen. In Abstimmung mit der Steuerungsgruppe hatte die Bewertungsgruppe Chemie die Zahl der Bewertungskriterien in den Dimensionen Nachwuchsförderung und Wissenstransfer reduziert.

|¹⁰ Die Bewertungsmatrix ist im Internet veröffentlicht unter:
www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/Bewertungsmatrix_ANAM.pdf

Bewertung der o. g. Kriterien erfolgte in der Anglistik und Amerikanistik vor dem Hintergrund der Überlegungen, dass bei qualitativen Informationen (Publikationen), die für die Bewertung des Kriteriums „Forschungsqualität“ in Anglistik und Amerikanistik maßgeblich waren, eine Effizienzberechnung und -bewertung nicht sinnvoll bzw. lediglich bei Publikationslisten möglich ist.

Die definierten Bewertungskriterien und Bewertungsaspekte wurden größtenteils durch qualitative Informationen operationalisiert: 26 von 32 der für die Bewertung einzelner Kriterien im Rahmen der Hauptdatenerhebung erhobenen Daten und Informationen waren qualitativer Natur. Damit sollte die geisteswissenschaftliche Forschungspraxis im Zuge der Bewertung angemessen berücksichtigt werden.

Um der Verfasstheit des Fachs Anglistik und Amerikanistik angemessen Rechnung zu tragen und um eine Binnendifferenzierung innerhalb der Einrichtungen zu ermöglichen, wurden alle Kriterien auf der Ebene der Teilbereiche erhoben und bewertet.

Die Datensammlung im Forschungsrating befindet sich im Spannungsfeld zwischen dem Anspruch, einerseits den Aufwand für die teilnehmenden Einrichtungen so gering wie möglich zu halten und andererseits ausreichend Informationen und Daten für eine valide und differenzierte Bewertung zu erhalten. |¹¹ Zudem ermöglicht ein umfangreiches Datenset weiterführende Sekundäranalysen. Die Prüfung daraufhin, welche Daten sich nicht in hinreichend guter Qualität erheben ließen oder aufgrund von Interpretationsschwierigkeiten kaum berücksichtigt wurden, gehört daher in besonderem Maße zu den Aufgaben der Pilotstudie. Solche Daten sollten künftig nicht mehr einbezogen werden. Bei einer Verstetigung des Verfahrens ist die Balance zwischen den verschiedenen berechtigten Anforderungen auszutarieren. |¹² Konkrete Hinweise zu den einzelnen Kriterien werden im Folgenden genannt.

Kriterium I „Forschungsqualität“

Der Forschungsqualität kommt bei der Bewertung von Forschungsleistungen eine besondere Bedeutung zu. Im Rahmen dieses Kriteriums wurde neben der

|¹¹ Vgl. Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011.

|¹² Derzeit bereitet der Wissenschaftsrat Empfehlungen zu einem Kerndatensatz Forschung vor mit dem Ziel, den Aufwand der Bereitstellung der Daten zu Zwecken der Forschungsbewertung an Hochschulen und außeruniversitären Forschungseinrichtungen durch einheitliche Vorhaltung und kontinuierliche Pflege klar definierter Daten zu reduzieren. Die Empfehlungen werden auch Aussagen zum Impelmentierungsprozess eines solchen Kerndatensatzes beinhalten. Der Wissenschaftsrat wird die Empfehlungen im Januar 2013 beraten.

„Qualität des Outputs“ auch die „Quantität des Outputs“ berücksichtigt. Die erhobenen Daten und Informationen lieferten insgesamt eine konsistente und somit aussagekräftige Datengrundlage und waren zur Bewertung der Forschungsqualität gut geeignet.

5 Anmerkungen zu einzelnen Daten:

– *Eingereichte exemplarische Publikationen und Publikationsauszüge*: Die Einreichung und die Lektüre von drei exemplarischen Publikationen pro Professur, die für die Bewertung der Forschungsqualität zentral war, hat sich bewährt. Insbesondere die Möglichkeit zur Berücksichtigung einer Publikation des wissenschaftlichen Nachwuchses war für den Einblick in die Gesamtleistung der Professur günstig und sollte auch künftig beibehalten werden.

Dabei erwies sich die Einschätzung von *Publikationsauszügen*, insbesondere aus Monographien, immer dann als schwierig, wenn die Gutachterinnen und Gutachter die Gesamtpublikation nicht kannten, da die eventuellen Monita für den eingereichten Ausschnitt nicht zwangsläufig für die gesamte Veröffentlichung zutreffen. Wissenschaftlerinnen und Wissenschaftlern sollte daher in Zukunft die Möglichkeit eingeräumt werden, gegebenenfalls die gesamte Monographie einzureichen. Um den ohnehin enormen Lektüre- und Bewertungsaufwand dennoch handhabbar zu halten, sollten die Autorinnen und Autoren gebeten werden, bestimmte Kapitel im Umfang von ca. 50 Seiten zu kennzeichnen, die für die jeweilige Veröffentlichung zentral und im Rahmen der Bewertung vordergründig zu berücksichtigen sind.

Als wenig aussagekräftig für die Bewertung der Forschungsqualität erwiesen sich eingereichte Handbuchartikel, Übersichtsdarstellungen sowie Einleitungen zu Sammelbänden. Für zukünftige Forschungsratings wird empfohlen, diese Veröffentlichungen bei der Erhebung der exemplarischen Publikationen explizit auszuschließen.

Die für die Bewertung der Veröffentlichungen gewählten Kriterien, nämlich Bedeutung, Innovationsgrad, Originalität, Aktualität, Ausstrahlung (national und international), forschungsmethodische Qualität sowie Breite und Einfluss der Fragestellung auf das eigene Forschungsfeld und andere Disziplinen, haben sich bewährt.

Publikationslisten: Die qualitative Begutachtung der Publikationslisten ergänzte die Lektüre von exemplarisch eingereichten Schriften gut. Eingeschränkt war die Verwendung dieser Informationen durch die von den Teilbereichen teilweise inkorrekt vorgenommene Zuordnung zum jeweiligen Publikationstyp (Monographien, Wissenschaftliche Editionen und Anthologien, Sammelbände, Zeitschriften- und Sammelbandaufsätze und Rezensionen). So wurden beispielsweise vielfach bei dem Publikationstyp „Monographie“ Angaben zu Lehrbüchern oder Diskussionspapieren gemacht. Einige Listen enthielten An-

gaben zu nicht wissenschaftlichen Publikationen. In manchen Listen wurden einzelne Angaben mehrfach gemacht. Die Kontrolle und Korrektur dieser Unstimmigkeiten war im Rahmen der Bewertung mit einem nicht unerheblichen Aufwand verbunden. Bei einer wiederholten Durchführung des Forschungsrating in der Anglistik und Amerikanistik sollten daher die von den Einrichtungen vorgenommenen Einträge frühzeitig einer fachlichen Plausibilitätsprüfung durch die Gutachterinnen und Gutachter unterzogen werden. Die Dubletten- sowie Plausibilitätsprüfung der Formalien sollten weiterhin durch die Geschäftsstelle des Wissenschaftsrats ausgeführt werden.

10 In diesem Rahmen könnten Einrichtungen eine Rückmeldung zu Mehrfachnennungen, Lehrwerken, nicht wissenschaftlichen Publikationen bzw. zu den für den Teilbereich nicht relevante Publikationen, die vor der Bewertung umsortiert bzw. aussortiert werden sollten, gegeben werden. Hilfreich wären zudem trennschärfere Vorgaben und Definitionen der einzelnen Publikationstypen. Dies gilt insbesondere zum einen mit Blick auf die Unterscheidung
15 zwischen wissenschaftlichen und nicht wissenschaftlichen Publikationen und zum anderen zwischen nicht wissenschaftlichen Publikationen, die gleichwohl dem wissenschaftlichen Diskurs zuzuordnen sind und solchen, die eher dem Transfer in andere Bereiche (Öffentlichkeit, Berufsfeld) dienen.

20 Die nach vorgegebenen Publikationstypen differenzierte Abfrage der Publikationen erwies sich gleichwohl weitestgehend als praktikabel und bedarf nur weniger Anpassungen. So haben viele Wissenschaftlerinnen und Wissenschaftler Lexikon- und Enzyklopädieartikel vielfach dem Publikationstyp „Zeitschriften- und Sammelbandaufsätze“ zugeordnet. Sinnvoll wäre hier die
25 Definition einer entsprechenden eigenständigen Kategorie. Um eine angemessene Bewertung der einzelnen Teilbereiche zu gewährleisten, sollten die Publikationstypen darüber hinaus teilbereichsspezifisch angepasst werden. So war die gängige Aufteilung der Publikationstypen für den Teilbereich Fachdidaktik Englisch teilweise problematisch, da nicht alle fachdidaktischen Publikationen diesen Typen zugeordnet werden können (z. B. Lehrmaterialien, elektronische Lernangebote, Curricula). Für den Teilbereich Englische Sprachwissenschaft erwies sich hingegen die Kategorie „Anthologien, Editionen“ als verzichtbar.

35 Die Erhebung und qualitative Begutachtung der Publikationslisten war hilfreich für die Beurteilung der Breite der Publikationsleistungen der jeweiligen Teilbereiche und sollte daher auch bei einer Wiederholung beibehalten werden. Quantitative Informationen (Anzahl der Publikationen differenziert nach Publikationstypen) boten dabei eine gute Ergänzung und rundeten das Bild ab.

Abweichend von Empfehlungen der Steuerungsgruppe zur vergleichenden Forschungsbewertung in den Geisteswissenschaften wurde „Reputation“ nicht als Bestandteil des Kriteriums „Forschungsqualität“ sondern als eigenständiges Kriterium bewertet. Die zu diesem Kriterium von den Einrichtungen vorgenommenen Einträge waren sowohl im Hinblick auf die Qualität als auch auf die Quantität sehr heterogen, wodurch die Bewertung dieses Kriteriums erschwert war.

Anmerkungen zu ausgewählten Daten:

10 _ *Wissenschaftliche Preise und Auszeichnungen*: Die hierzu von den Einrichtungen – sofern überhaupt – vorgenommen Einträge waren im Hinblick auf ihre Relevanz für die Bewertung von Reputation sehr heterogen. Zugleich ist fraglich, ob trennschärfere Vorgaben Abhilfe geschaffen hätten.

15 Ausgehend von der Annahme, dass die Forschenden Preisverleihungen nicht aktiv beeinflussen können, war eine Nullmeldung bei wissenschaftlichen Preisen und Auszeichnungen im Vergleich zu anderen Daten im Rahmen der Bewertung weniger bedeutsam. Sofern Forschungspreise und Auszeichnungen genannt waren, wurden diese bei der Bewertung berücksichtigt.

20 _ *Gastwissenschaftlerinnen und Gastwissenschaftler*: Auch die hierzu vorgenommen Einträge waren sehr heterogen. Entgegen einer sehr präzise gefassten Definition des Begriffs wurden hier teilweise auch Gastvorträge aufgelistet.

Aus diesen Gründen war der Bewertungsaspekt „Anerkennung“ in sehr vielen Fällen nicht bewertbar und das Kriterium II „Reputation“ musste ausschließlich auf Grundlage der Angaben zum Bewertungsaspekt „*professional activities*“ beurteilt werden, welches sich trotz teilweiser Mehrfachnennungen einerseits und Auslassungen andererseits als aussagekräftig erwies. Der Bewertungsaspekt „Anerkennung“ hat sich somit nicht bewährt, künftig sollte auf diesen Bewertungsaspekt verzichtet werden. Die für diesen Aspekt relevanten Angaben könnten beispielsweise als ergänzende Angaben zur Reputation in Form eines Kommentars vorgenommen werden.

Die Bewertung des Kriteriums „Reputation“ als eigenständiges Kriterium hat sich bewährt und sollte auch bei künftigen Verfahren beibehalten werden. Um die Datenqualität zu verbessern, sollte im Zuge der Datenerhebung die Definition des Kriteriums und der Bewertungsaspekte genauer spezifiziert werden.

Mit dem Bewertungskriterium „Forschungsermöglichung“ sollten wissenschaftsimmanente Tätigkeiten erfasst werden, die das Erbringen von Forschungsleistungen überhaupt erst ermöglichen. |¹³ Die zur Bewertung dieses Kriteriums ausgewählten Bewertungsaspekte und Daten haben sich grundsätzlich bewährt. Insbesondere die quantitativen Daten und Kennzahlen trugen zur Vereinfachung und Transparenz der Bewertungen bei.

Anmerkungen zu ausgewählten Daten:

10 _ *Drittmittelprojekte und verausgabte Drittmittel*: Die nach Mittelgebern und Jahren aufgeschlüsselte Abfrage der verausgabten Drittmittel für die einzelnen Teilbereiche war weitgehend unproblematisch. Lediglich eine Einrichtung konnte keine nach Teilbereichen differenzierten Angaben machen und bei einigen Einrichtungen bestand Korrekturbedarf, da trotz präziser Definition |¹⁴ der beiden Daten Angaben gemacht wurden, die der Definition nicht entsprachen
15 (z. B. Reisemittel und Tagungsmittel). Optimierungsmöglichkeit bei der Erhebung der Informationen zu Drittmittelaktivitäten wird in der Anpassung des Erhebungsprinzips für die Drittmittelprojekte und die verausgabten Drittmittel gesehen. Da einerseits die im Erhebungszeitraum bewilligten Drittmittelprojekte und andererseits die in den einzelnen Jahren des Erhebungszeitraums verausgabten Drittmittel erhoben wurden, bestand nicht immer eine
20 Verbindung zwischen den beiden Informationen, die Beurteilung war dadurch erschwert. Um die angegebenen Drittmittelprojekte eindeutig identifizieren zu können, sollte künftig zusätzlich das Förderkennzeichen erhoben werden.

25 _ *Laufende Promotionen*: Die von den Einrichtungen vorgelegten Listen der laufenden Promotionen erwiesen sich als nicht aussagekräftig. Die Bewertung dieser Informationen war schwierig, da der Abschlusserfolg der gelisteten Dissertationen nicht mit Sicherheit vorhergesagt werden kann. Diesem Datum wurde im Rahmen des Bewertungsprozesses entsprechend eine geringe
30 Bedeutung beigemessen.

_ *Abgeschlossene Promotionen*: Die Bewertungsgruppe hatte sich zu Beginn des Verfahrens gegen eine Beurteilung der Erfolge in der Nachwuchsförderung auf Grundlage der Zahl der abgeschlossenen Promotionen ausgesprochen, da

|¹³ Vgl. Wissenschaftsrat: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, Köln 2010, S. 28.

|¹⁴ So war die Angabe der Mittel der personenbezogenen Förderung (Stipendien), Reisekosten-, Tagungs- und Druckkostenzuschüsse explizit von der Erhebung ausgeschlossen. (Vgl. Glossar, Stichwort Drittmittel: <http://www.wissenschaftsrat.de/download/archiv/Glossar.pdf>)

diese zwar Auskunft über die Quantität nicht jedoch über die Qualität der Nachwuchsförderung geben. Diese Vorgehensweise hat sich weitgehend bewährt. Um eine präzisere Beurteilung des Erfolgs der Nachwuchsförderung zu ermöglichen, sollten diese Informationen dennoch künftig um quantitative Angaben zu abgeschlossenen Promotionen ergänzt werden. Für die Beurteilung der Erfolge der Nachwuchsförderung im Rahmen des Bewertungsprozesses waren die erhobenen qualitativen Informationen zu abgeschlossenen Promotionen (Name der Promovendin/des Promovenden, Name der Erstbetreuerin/des Erstbetreuers, Titel und Jahr) wichtiger als Informationen zu laufenden Promotionen.

– *Konferenzen und Tagungen*: Die Informationen zu veranstalteten Konferenzen und Tagungen sind künftig um Angaben zu Dauer und Teilnehmerzahl zu ergänzen.

– *Netzwerke und Verbände*: Eine angemessene Bewertung von Angaben zu Netzwerken und Forschungsverbänden, an denen die gemeldeten Wissenschaftlerinnen und Wissenschaftler maßgeblich beteiligt waren, war aufgrund stark heterogener Einträge, deren Bedeutung erheblich variierte, schwierig. Teilweise wurden neben bedeutenden nationalen und internationalen Netzwerken, Forschungsverbänden und Zentren auch wenig bedeutende und informelle Verbände u. ä. gelistet. Künftig sollte dieses Datum trennschärfer beschrieben werden.

Kriterium IV „Transfer an außeruniversitäre Adressaten“

Im Rahmen dieses Kriteriums wurde der Beitrag der Einrichtungen zum forschungsbasierten Wissenstransfer bewertet, aufgeschlüsselt nach Personaltransfer und Wissenstransfer. Von den einzelnen Einrichtungen wurde diesem Kriterium unterschiedliche Bedeutung beigemessen, so dass die Qualität der hierzu erfolgten Einträge stark variiert. Bei einigen Einrichtungen bestanden Zweifel, welcher Art Angaben an dieser Stelle gemacht werden sollten. Statt der wichtigsten und bedeutendsten Maßnahmen und Aktivitäten hat ein Teil der Einrichtungen offenbar vollständige Listen eingereicht, andere hingegen haben nur lückenhafte oder überhaupt keine Angaben vorgenommen. Dadurch war eine angemessene Bewertung dieses Kriteriums erschwert. Auch war die von den Einrichtungen getroffene Unterscheidung zwischen wissenschaftlichen Aktivitäten/Tätigkeiten und solchen, die eher dem Bereich Transfer zuzuordnen sind, für die Gutachterinnen und Gutachter nicht immer nachvollziehbar.

Trotz der oben beschriebenen Schwierigkeiten und mit Blick auf die steigende Bedeutung des Transfers von Forschungsergebnissen, sollte auf die Erfassung und Bewertung der Transferleistungen, insbesondere solcher an die außeruniversitäre Adressaten auch künftig nicht verzichtet werden. Die Aufteilung in die Bewertungsaspekte „Personaltransfer“ und „Wissenstransfer“ hat sich nicht

bewährt, da sich diese in der Ausfüllpraxis nicht immer widerspiegelten. Bei künftigen Erhebungen sollte dieses Kriterium daher durch trennschärfere Bewertungsaspekte und präzisere Erhebungsanweisungen definiert werden.

5 Zugleich sollten die einzelnen Angaben ausführlicher beschrieben und durch Beispiele präzisiert werden. Darüber hinaus sollten die einzelnen Informationen differenzierter erhoben werden (z. B. Dauer und Typen von Fortbildungsveranstaltungen). Schließlich zeigte sich, dass die Bedeutung von Transferleistungen zwischen einzelnen Teilbereichen variiert, so dass auch diesbezüglich eine entsprechende Anpassung erforderlich ist.

10 *Rahmeninformationen*

Die Rahmeninformationen dienten im Rahmen der Bewertung der Einordnung aller übrigen Daten. Die erhobenen Rahmeninformationen zu Einrichtungen und Teilbereichen erwiesen sich als ausgesprochen aussagekräftig und hilfreich. Die Möglichkeit der Selbstbeschreibung sowie der Beschreibung der Forschungsschwerpunkte erlaubte eine Kontextualisierung der Forschungsaktivitäten, insbesondere der Publikationen. Die Angaben zur Lehr- und Prüfungsauslastung einerseits sowie zur Personalausstattung andererseits haben zum Teil Hinweise für die Erklärung von fehlenden Aktivitäten in anderen Bereichen geliefert. Um den Aufwand der Bewertung nicht zu erhöhen sollten die Selbstbeschreibungen auch weiterhin auf eine angemessene Länge begrenzt erhoben werden.

Ausgesprochen nützlich waren insbesondere die Angaben zu Vakanzen. Um diese Informationen künftig noch systematischer im Rahmen des Bewertungsprozesses berücksichtigen sowie in die Veröffentlichung der Ergebnisse integrieren zu können, ist eine Standardisierung ihrer Erhebung erforderlich.

Trotz des ausgesprochen hohen Nutzwertes der Rahmeninformationen für die Einordnung der übrigen Daten erwiesen sich diese als nicht ausreichend. Im Sinne einer objektiveren Berücksichtigung der vorhandenen Ressourcen, sollte auf eine gesonderte Berechnung und Bewertung der Effizienz künftig nicht verzichtet werden.

1.3 Bedingungen der Teilnahme

Am Forschungsrating Anglistik und Amerikanistik konnten alle Einrichtungen teilnehmen, die in mindestens einem der definierten Teilbereiche forschungsaktiv waren. Zudem sollten sich nur solche Einrichtungen beteiligen, die mindestens die Hälfte des Erhebungszeitraums (01.01.2004 – 31.12.2010) bestanden haben. Untergrenzen im Hinblick auf die personelle Ausstattung wurden nicht formuliert.

Wie in der ersten Pilotstudie, so war auch in der Anglistik und Amerikanistik die Erfassungsquote sehr hoch. Mit 358 Professuren zum Stichtag erreicht das Fach 94 Prozent der vom Statistischen Bundesamt dem Lehr- und Forschungsbereich „Anglistik, Amerikanistik“ zugeordneten Professuren (379 im Jahr 2010). |¹⁵

Die gemeldeten Teilbereiche in der Anglistik/Amerikanistik verfügen häufig über nur eine Professur: in 87 Teilbereichen (45,55 Prozent der Teilbereiche) wurde nur eine Professur gemeldet. Die Großzahl dieser Teilbereiche sind Amerika-
 10 studien und der Fachdidaktik Englisch zugeordnet, ein kleinerer Teil findet sich in der Anglistik: Literatur- und Kulturwissenschaft. Diese Besonderheit war sowohl aus Sicht des Datenschutzes als auch mit Blick auf das Ziel des Verfahrens, nämlich einrichtungsbezogene Bewertungen vorzunehmen, nicht unproblematisch. Die Bewertungsgruppe war gleichwohl der Überzeugung, dass auch diese
 15 Teilbereiche bewertbar sind, da Forschungsleistungen in der Anglistik und Amerikanistik üblicherweise nicht nur von Professorinnen und Professoren sondern in einem erheblichen Maße auch von wissenschaftlichen Mitarbeiterinnen und Mitarbeitern erbracht werden, die die Teilbereiche mitkonstituieren. Aus datenschutzrechtlichen Gründen wurden jedoch Bewertungsergebnisse für
 20 Teilbereiche, in denen weniger als drei Wissenschaftlerinnen und Wissenschaftler |¹⁶ gemeldet wurden nicht veröffentlicht. Lediglich die betreffenden Einrichtungen und das jeweilige Sitzland wurden über das Ergebnis der Bewertung informiert. Betroffen sind 35 Teilbereiche an 27 Einrichtungen (16,75 Prozent der Teilbereiche).

25 In insgesamt 16 Teilbereichen wurde ursprünglich keine Professur gemeldet. Um eine faire Bewertung der Teilbereiche zu gewährleisten, wurden die betroffenen Einrichtungen gebeten zu prüfen, ob die für diese Teilbereiche gemeldeten Wissenschaftlerinnen und Wissenschaftler einen Forschungsauftrag haben. Bei positiver Rückmeldung wurde den Einrichtungen vorgeschlagen, wenn möglich, eine Zuordnung dieser Wissenschaftlerinnen und
 30 Wissenschaftler zu einem anderen Teilbereich vorzunehmen. Wichtig war diese Vorgehensweise nicht zuletzt mit Blick auf die Einreichung von exemplarischen Publikationen, deren Begutachtung für die Bewertung des Kriteriums I „Forschungsqualität“ maßgeblich war und die nur für Professuren eingereicht
 35 werden sollten. Die Meldung von 14 Teilbereichen wurde vor diesem Hinter-

| ¹⁵ Statistisches Bundesamt: Bildung und Kultur, Personal an Hochschulen, Fachserie 11, Reihe 4.4, Wiesbaden 2010, S. 94.

| ¹⁶ Hierzu zählen Professorinnen und Professoren, weitere habilitierte Wissenschaftlerinnen und Wissenschaftler sowie wissenschaftliche Mitarbeiterinnen und Mitarbeiter.

grund zurückgezogen und für die gemeldeten Wissenschaftlerinnen und Wissenschaftler wurde zum Teil eine Neuordnung vorgenommen. Zwei Teilbereiche haben letztlich ohne Professur am Forschungsrating teilgenommen. Allen Einrichtungen wurde im Rahmen der Hauptdatenerhebung zugleich die Möglichkeit eingeräumt, auf zum Stichtag 31.12.2010 bestehende Vakanzen hinzuweisen sowie anzugeben, ob aus diesem Grund Teilbereiche für die Teilnahme nicht gemeldet werden konnten.

Im Unterschied zu den anderen drei Pilotstudien in Chemie, Soziologie und Elektrotechnik und Informationstechnik haben am Forschungsrating Anglistik und Amerikanistik keine außeruniversitären Forschungseinrichtungen teilgenommen. Begrüßenswert war die erstmalige Teilnahme der Pädagogischen Hochschulen. Für die Zuordnung des wissenschaftlichen Personals zu definierten Teilbereichen und für die Kontextualisierung der erhobenen Daten im Rahmen der Bewertung ist die Besonderheit relevant, dass die dort forschenden und lehrenden Wissenschaftlerinnen und Wissenschaftler stets gleichzeitig eine Fachwissenschaft und die Fachdidaktik vertreten.

1.4 Vorbereitung der Erhebung

Entwicklung von Fragebögen

Alle für die Bewertung relevanten Daten mussten bei den teilnehmenden Einrichtungen erfragt werden. Auf Basis der „Bewertungsmatrix“ entwickelte die Bewertungsgruppe hierzu Fragebögen und ein Glossar, in dem alle wichtigen Begriffe erläutert und zum Teil durch Beispiele präzisiert wurden. Entsprechend den Empfehlungen der Steuerungsgruppe zur Vereinfachung der Datenerhebung wurden in Zusammenarbeit mit dem Umfragezentrum Bonn – Prof. Rudinger GmbH – Gesellschaft für empirische Sozialforschung und Evaluation (UZ) Online-Fragebögen entwickelt. Während die technische Realisierung des Erhebungsinstruments und die Datensammlung dem UZ übertragen wurden, lag die inhaltliche Vorbereitung der Erhebung sowie die Kontrolle und Aufbereitung der Daten bei der Geschäftsstelle des Wissenschaftsrats. Diese organisierte das Verfahren in Abstimmung mit den Gutachterinnen und Gutachtern und den Einrichtungen. Auch die Erhebung der exemplarischen Publikationen lag in der Verantwortung der Geschäftsstelle.

Obwohl alle Daten auf der Ebene der Teilbereiche bewertet wurden, wurden einige Rahmeninformationen auf der Ebene der Einrichtung, also für die Gesamtheit der Teilbereiche, und die Publikationslisten auf der Ebene der einzelnen Wissenschaftlerinnen und Wissenschaftler erfasst. Entsprechend setzte sich das Erhebungsinstrument aus drei Teilen – Fragebogen Einrichtung übergreifend, Fragebogen Teilbereich sowie Fragebogen Publikationen – zusammen. Pro Einrichtung wurden so viele Teilbereichsfragebögen zur Verfügung gestellt,

wie von der Einrichtung Teilbereiche gemeldet wurden. Die Fragebögen zur Erfassung der Publikationen waren personifiziert und orientierten sich an den für die Erhebung gemeldeten Wissenschaftlerinnen und Wissenschaftlern.

5 Mitte Oktober erhielten die Fachkoordinatorinnen und Fachkoordinatoren der teilnehmenden Einrichtungen per E-Mail die Zugangsdaten zum Online-Fragebogen vom UZ. Die Erteilung weiterer Zugriffsrechte („Rechtmanagement“) lag ausschließlich beim jeweiligen Fachkoordinator, so dass mehrere Nutzer innerhalb der Einrichtungen den Fragebogen bearbeiten konnten. Technische Probleme mit dem Fragebogen betrafen den Zugang und die Weiterleitung der Freigaben sowie den Abschluss des Fragebogens. Insgesamt konnten
10 die technischen Probleme seitens des UZ jedoch sehr schnell behoben werden.

Das neue Online-Format, das sich bereits im Rahmen des Forschungsratings Elektrotechnik und Informationstechnik bewährt hat, hat auch hier das Erhebungsverfahren deutlich vereinfacht und zu hoher Transparenz im Verfahren
15 beigetragen. Das webgestützte Erhebungsformat stieß auf breite Zustimmung sowohl bei der Bewertungsgruppe als auch bei den teilnehmenden Einrichtungen, die insbesondere die Möglichkeiten der Weiterleitung einzelner Fragekomplexe bzw. das damit einhergehende Rechtmanagement begrüßten. Die Online-Erhebung sollte in künftigen Forschungsratings beibehalten werden.

20 *Informationsveranstaltungen für die Fachkoordinatorinnen und Fachkoordinatoren*

Die Datenerhebung wurde in den teilnehmenden Einrichtungen federführend koordiniert durch Wissenschaftlerinnen und Wissenschaftler (Fachkoordinatorinnen und Fachkoordinatoren) aus dem Fach Anglistik und Amerikanistik, die von den jeweiligen Einrichtungsleitungen eingesetzt wurden. Damit sollte sichergestellt werden, dass die inhaltliche Kontrolle in der Hand der Wissenschaft lag. Um die Effizienz des Erhebungsprozesses zu steigern, benannten
25 Einrichtungen außerdem jeweils eine Ansprechpartnerin bzw. einen Ansprechpartner aus der Verwaltung zur Unterstützung ihrer Fachkoordinatorin bzw. ihres Fachkoordinators. Wie bereits im Forschungsrating Elektrotechnik und Informationstechnik hat die Geschäftsstelle des Wissenschaftsrats zusammen
30 mit Mitgliedern der Bewertungsgruppe und Vertretern des UZ im September 2011 zwei Informationsveranstaltungen für die Fachkoordinatorinnen und Fachkoordinatoren durchgeführt. Soweit möglich, gingen die dort eingebrachten Vorschläge der Fachkoordinatorinnen und Fachkoordinatoren in die Fertigstellung der Fragebögen ein. So wurden beispielsweise Frage nach Vakanzen
35 zum Stichtag 31.12.2010 eingefügt und die zu erhebenden Publikationstypen um „Rezensionen“ ergänzt. Die frühzeitige Information der Einrichtungen hat sich erneut bewährt und sollte künftig beibehalten werden.

II.1 Prinzipien der Datenerhebung

Für die Bewertung der Forschungsleistungen in der Anglistik und Amerikanistik wurde ein siebenjähriger Erhebungszeitraum (01.01.2004 – 31.12.2010) festgelegt und somit der in den bisherigen Pilotstudien verwendete fünfjährige Erhebungszeitraum ausgedehnt. Damit folgte die Bewertungsgruppe den Empfehlungen der Steuerungsgruppe zur vergleichenden Forschungsbewertung in den Geisteswissenschaften. |¹⁷ Ein ausreichend langer Bewertungszeitraum sollte der geisteswissenschaftlichen Forschungspraxis, die nach wie vor durch Individualforschung und daher durch langfristig angelegte Forschungsvorhaben gekennzeichnet ist, entsprechen. Die gewählte Dauer des Erhebungszeitraums hat sich für das Fach „Anglistik und Amerikanistik“ bewährt. Wichtig für die Interpretation der Ergebnisse ist der Umstand, dass gerade in diesen Zeitraum an einer Großzahl der teilnehmenden Einrichtungen die Umstellung auf die modularisierte Studienstruktur im Fach Anglistik und Amerikanistik fiel und damit viele Kapazitäten gebunden waren.

Die Datenerhebung in der Anglistik und Amerikanistik erfolgte nach dem *current-potential*-Prinzip (Leistung der zum Stichtag 31.12.2012 an der betreffenden Einrichtung beschäftigten Personen für den zurückliegenden 7-Jahres-Zeitraum). Da jedoch nicht alle erforderlichen Daten an den Einrichtungen stichtagbezogen vorliegen, wurde ein Teil der Informationen nach dem *work-done-at* Prinzip erhoben. (Leistung aller an der betreffenden Institution im 7-Jahres-Zeitraum vom 01.01.2004 bis 31.12.2010 beschäftigten Wissenschaftlerinnen und Wissenschaftler). Somit lag bei der Datenerhebung ein so genannter „hybrider“ Ansatz von *current-potential* und *work-done-at* zugrunde. |¹⁸ Mit dieser Vorgehensweise schließt das Forschungsrating Anglistik und Amerikanistik an die Pilotstudie Elektrotechnik und Informationstechnik an. Die Bewertungsgruppe Elektrotechnik und Informationstechnik kam nach Abwägen der Argumente und der Auswertung der Erfahrungen zu dem Schluss, dass ein hybrider Ansatz den bestmöglichen Kompromiss darstellt. Das *current-potential*-Prinzip wurde in der Anglistik und Amerikanistik bei der Erhebung der Studienfälle (einschl. Lehramtsstudierende), Vakanzen, Personaldaten, Publikationen, der Forschungspreise und wissenschaftlicher Auszeichnungen, Akademischer Äm-

| ¹⁷ Vgl. Wissenschaftsrat: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, Köln 2010, S. 24 f.

| ¹⁸ Die Datenerhebung wurde in der Pilotstudie Forschungsrating für das Fach Chemie nach dem Prinzip *work-done-at* durchgeführt, im Fach Soziologie nach dem *current-potential*-Prinzip und im Forschungsrating Elektrotechnik und Informationstechnik wurde ein wie oben beschriebener „hybrider Ansatz“ gewählt.

ter, Akademiemitgliedschaften, laufende Promotionen, Promotionsprogramme, Stipendien, Gutachtertätigkeiten, Editorial Board Mitgliedschaften, Infrastrukturen, Netzwerke und Verbände sowie Konferenzen und Tagungen angewandt. Bei der Erhebung der Gastwissenschaftlerinnen und Gastwissenschaftler, der bewilligten Drittmittelprojekte, verausgabter Drittmittel, abgeschlossener Promotionen, abgeschlossener Habilitationen, Auszeichnungen, Preise und Erstrufe, Weiterbildungsangebote, Kooperationen/Ämter, Öffentlichkeitsarbeit und Medien sowie Forschungsprodukte wurde nach dem *work-done-at*-Prinzip verfahren. Beide Ansätze sind mit unterschiedlichen Vor- und Nachteilen verbunden. Die Erhebungen nach dem „*work-done-at*“-Prinzip sind unempfindlicher für Fluktuationen der Wissenschaftlerinnen und Wissenschaftler, da hierbei alle Forschungsleistungen berücksichtigt werden, die an der jeweiligen Einrichtung erbracht wurden, unabhängig davon, wo die Personen, die sie erbracht haben zum Stichtag beschäftigt sind. Zugleich liefern jedoch auch die Erhebungen nach diesem Prinzip keine zuverlässigen Daten für solche Teilbereiche, an denen langjährige Vakanzen im Erhebungszeitraum bestanden. Die Bewertungen auf Basis der nach dem *work-done-at*-Prinzip erhobenen Daten sind weniger personengebunden und erlauben somit eher Aussagen über den gesamten Teilbereich. Da viele Daten an den Einrichtungen jedoch stichtagbezogen vorgehalten werden, ist die Erhebung nach diesem Prinzip häufig erschwert, mit Unschärfen bei der Zuordnung von Forschungsleistungen zu Einrichtungen behaftet und für einige Daten derzeit häufig nicht möglich (z. B. Forschungspreise und Auszeichnungen). Zudem erlauben die so erhobenen Daten nur eingeschränkt prospektive Aussagen. Demgegenüber erfasst eine Erhebung nach dem *current-potential*-Prinzip die zum Stichtag aktuelle Situation und Leistungsfähigkeit einer Einrichtung und ermöglicht somit Handlungsempfehlungen für die Zukunft. Dieses Erhebungsprinzip entspricht somit der Intention des Forschungsratings zukunftsgerichtete Entwicklungspotentiale zu erfassen. Zudem gestaltet sich die Datenbereitstellung für die Einrichtungen einfacher, da alle relevanten Daten bei den zum Stichtag beschäftigten Wissenschaftlerinnen und Wissenschaftlern erfragt werden können. Diese Vorgehensweise kann jedoch immer dann zu Verzerrungen führen, wenn im Erhebungszeitraum hohe Personalfluktuation herrschte oder Stellen zum Stichtag nicht besetzt waren. Um diesem Nachteil zu begegnen wurden im Rahmen der Erhebung auch Vakanzen zum Stichtag sowie Angaben zu Fluktuationen erfragt. Eine adäquate Abbildung von Generationenwechseln stellt im Rahmen von vergleichenden Bewertungsverfahren eine besondere Herausforderung dar. Eine besondere Schwierigkeit, die sich am stärksten in den Kriterien „Reputation“ und „Forschungsermöglichung“ niederschlägt, ist die Erfassung von Forschungsleistungen der jungen Forschenden und der Emeriti.

Die relativ lange Projektdauer des Forschungsratings Anglistik und Amerikanistik, die in erster Linie dem *peer-review*-Verfahren geschuldet ist, bedingte einen

Stichtag (31.12.2010) der fast zehn Monate vor der Datenerhebung und knapp zwei Jahre vor der Ergebnisveröffentlichung lag. Sofern das Forschungsrating auch künftig in Form eines *informed-peer-review* stattfinden wird, wird sich die Dauer einzelner Ratings auch bei Wiederholung nicht wesentlich verkürzen lassen. Bei der Wahl des Stichtags wurde zudem berücksichtigt, dass viele Daten an den Einrichtungen immer nur zum 31.12. eines Jahres vorliegen. Damit konnten jedoch viele aktuelle Publikationen sowie andere Forschungsaktivitäten im Zuge dieser Studie nicht berücksichtigt werden. Bei einer wiederholten Durchführung des Forschungsratings würden die nach dem 31.12.2010 erbrachten Leistungen jedoch erfasst werden und in die Bewertungen einfließen. Davon unabhängig spricht sich die Bewertungsgruppe für einen möglichst aktuellen Stichtag aus. Um dies zu gewährleisten, sollten die Projektplanung und -verlauf derart terminiert werden, dass die Datenerhebung möglichst nah zum Stichtag erfolgen kann. Zudem könnte für einzelne Daten (z. B. Publikationen) ein abweichender bzw. aktuellerer Stichtag gewählt werden.

II.2 Organisation und Ablauf

Die Datenerhebung erfolgte in drei Schritten: 1. Personenerhebung, 2. Publikationserhebung sowie 3. Hauptdatenerhebung. Zu jedem Erhebungsschritt wurden die Fachkoordinatorinnen und Fachkoordinatoren durch eine personalisierte E-Mail eingeladen. Zu jeder einzelnen Phase der Erhebung erhielten die Fachkoordinatorinnen und Fachkoordinatoren entsprechende Handreichungen.

Personenerhebung

Zunächst wurden die Einrichtungen gebeten, ihre Forschungsaktivitäten den vier Teilbereichen zuzuordnen und für diese jeweils die Professorinnen und Professoren, weitere habilitierte Wissenschaftlerinnen und Wissenschaftler, wissenschaftliche Mitarbeiterinnen und Mitarbeiter sowie promovierte Stipendiatinnen und Stipendiaten namentlich zu benennen, die zum Stichtag 31.12.2010 an der Einrichtung beschäftigt waren. |¹⁹ Diese Erhebung fand im Mai und Juni 2011 statt. Einzelne Abmeldungen, Nachmeldungen und Umwidmungen von Wissenschaftlerinnen und Wissenschaftlern fanden jedoch noch bis zur Hauptdatenerhebung im Oktober 2011 statt. Im Ergebnis wurden an 60 Einrichtungen (davon sechs Pädagogische Hochschulen) insgesamt 190 Teilbereiche |²⁰ gemeldet und diesen 1158 Wissenschaftlerinnen und Wissenschaftler,

| ¹⁹ Auch Beurlaubungen und Vertretungen konnten genannt werden, sofern die betreffenden Personen in einem Beschäftigungsverhältnis mit der Einrichtung standen.

| ²⁰ Im Rahmen der Hauptdatenerhebung zog eine Einrichtung und ein Teilbereich aus verschiedenen Gründen ihre Teilnahme zurück und sind in den genannten Zahlen nicht berücksichtigt.

davon 358 Professorinnen und Professoren zugeordnet. Die namentliche Erfassung diente insbesondere der Erhebung von Publikationslisten im Rahmen der Hauptdatenerhebung, die Namen der Professorinnen und Professoren wurden zudem im Rahmen der Erhebung von exemplarischen Publikationen eingesetzt. Promovierte Stipendiatinnen und Stipendiaten, insgesamt 20 Personen, wurden an nur wenigen Einrichtungen gemeldet. Diese Kategorie erwies sich somit als wenig relevant.

Im Rahmen dieser Erhebung erreichten die Geschäftsstelle viele Nachfragen, die in der Regel Zuordnungsfragen zu Teilbereichen und zu Personalkategorien sowie Handhabung von Beurlaubungen und von Vertretungen betrafen. Nach Abschluss der Erhebung wurden die Einträge der Einrichtungen auf Plausibilität sowie auf Doppel- und Mehrfachnennungen hin geprüft. Sofern eine Person mehr als einem Teilbereich zugeordnet oder von mehr als einer Einrichtung gemeldet wurde, erfolgte eine Rückfrage bei den betreffenden Einrichtungen. (vgl. I.3)

Publikationserhebung

Ende August 2011 wurden die Fachkoordinatorinnen und Fachkoordinatoren gebeten, bis Mitte Oktober 2011 für jede gemeldete Professur drei Publikationen bzw. Publikationsausschnitte à maximal 50 Seiten einzureichen (zuzüglich der Titelseite und des Inhaltsverzeichnisses bei Monographien und Sammelbänden). Die Veröffentlichungen sollten in dem Erhebungszeitraum 01.01.2004 – 31.12.2010 publiziert (*print* oder *online*) worden sein. Die Erfassung von exemplarischen Publikationen erfolgte für die folgenden Publikationskategorien: Monographien, Sammelbände/wissenschaftliche Editionen/Anthologien sowie Zeitschriften- und Sammelbandartikel. Um die Leistungen des wissenschaftlichen Nachwuchses zu berücksichtigen, konnte eine der Publikationen von einer Mitarbeiterin oder einem Mitarbeiter der Professur eingereicht werden, die zuvor namentlich gemeldet worden waren. Grundsätzlich konnten alle Publikationen unabhängig von der Teilbereichszuordnung der gemeldeten Wissenschaftlerinnen und Wissenschaftler eingereicht werden. Auf Publikationen bzw. Publikationsausschnitte, die nach Einschätzung der Fachkoordinatorinnen und Fachkoordinatoren nicht ausschließlich von den Gutachterinnen und Gutachtern des Teilbereichs begutachtet werden konnten, dem die Autorin bzw. der Autor zugeordnet wurde, sollte gesondert hingewiesen werden. Von dieser Möglichkeit haben jedoch nur wenige Einrichtungen Gebrauch gemacht. Die Publikationen bzw. Publikationsausschnitte sollten in digitaler Form auf einer DVD eingereicht werden. Sofern eine Ersatzpublikation eingereicht werden musste (s. u.), erfolgte dies in der Regel per E-Mail.

Nach dem Eingang der Publikationen bzw. Publikationsausschnitte wurden diese von der Geschäftsstelle in einer Datenbank erfasst und auf Vollständigkeit

sowie darauf hin geprüft, ob sie den oben beschriebenen Anforderungen genügen. Sofern das Jahr der Veröffentlichung aus der eingereichten Kopie nicht eindeutig hervorging, wurde es von der Geschäftsstelle recherchiert und auf der Publikation entsprechend vermerkt. Die Erhebung der exemplarischen Publikationen war insgesamt mit einer Vielzahl von Rückfragen an die Einrichtungen verbunden, die im Ergebnis oft eine Einreichung von Ersatzpublikationen oder Kürzungen bedeuteten. Die häufigsten Rückmeldungen und Nachfragen betrafen:

10 _ Erscheinungsjahr der eingereichten Veröffentlichung lag vor oder nach dem Erhebungszeitraum: Die Einrichtungen wurden um Einreichung von Ersatzpublikationen gebeten.

_ Länge der eingereichten Veröffentlichung: Die Einrichtungen wurden gebeten, eine gekürzte Fassung einzureichen oder mitzuteilen, welche Teile der Veröffentlichung zur Verfügung gestellt werden sollen.

15 _ Fehlen des Titelblatts und/oder des Inhaltsverzeichnisses: Die Einrichtungen wurden gebeten, die fehlenden Seiten nachzureichen.

_ zu viele Publikationen bzw. Publikationsausschnitte: Die Einrichtungen wurden gebeten mitzuteilen, welche der eingereichten Publikationen der Bewertung zugrunde gelegt werden sollen.

20 Nach Abschluss aller erforderlichen Nacherhebungen im November 2011, übermittelte die Geschäftsstelle die eingereichten exemplarischen Publikationen bzw. Publikationsausschnitte an die Gutachterinnen und Gutachter. Dabei erhielt jedes Mitglied der Bewertungsgruppe eine DVD mit Publikationen der ihr bzw. ihm zugeordneten Einrichtungen. Jede Publikation einer Professur war als eigenständige Datei gespeichert und der Einrichtung sowie der betreffenden
25 Professur durch den Dateinamen eindeutig zugeordnet. Zusätzlich erhielten die Gutachterinnen und Gutachter ein Inhaltsverzeichnis der übermittelten Publikationen mit Informationen zu Autorinnen/Autoren, Titel, Erscheinungsjahr sowie Zeitschrift/Verlag.

30 Trotz des hohen Aufwandes bei der Prüfung und Korrektur der Einreichungen, hat sich der Modus der Erhebung von exemplarischen Publikationen bewährt und sollte beibehalten werden. Im Rahmen künftiger Erhebungen sollte eine bessere Qualität der digitalen Kopien gefordert werden.

35 Die folgenden Abbildungen geben einen Überblick über die Struktur (Jahr und Publikationstyp) der eingereichten exemplarischen Publikationen bzw. Publikationsausschnitte differenziert nach Teilbereichen.

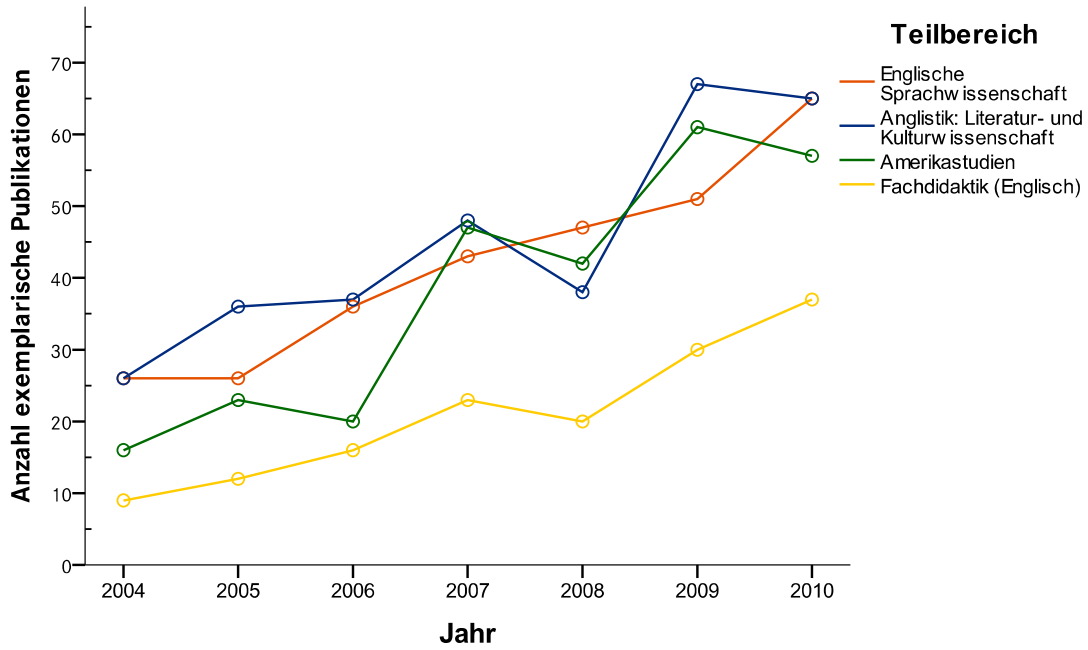


Abbildung 2: Anteil exemplarischer Publikationen nach Teilbereichen und Jahren

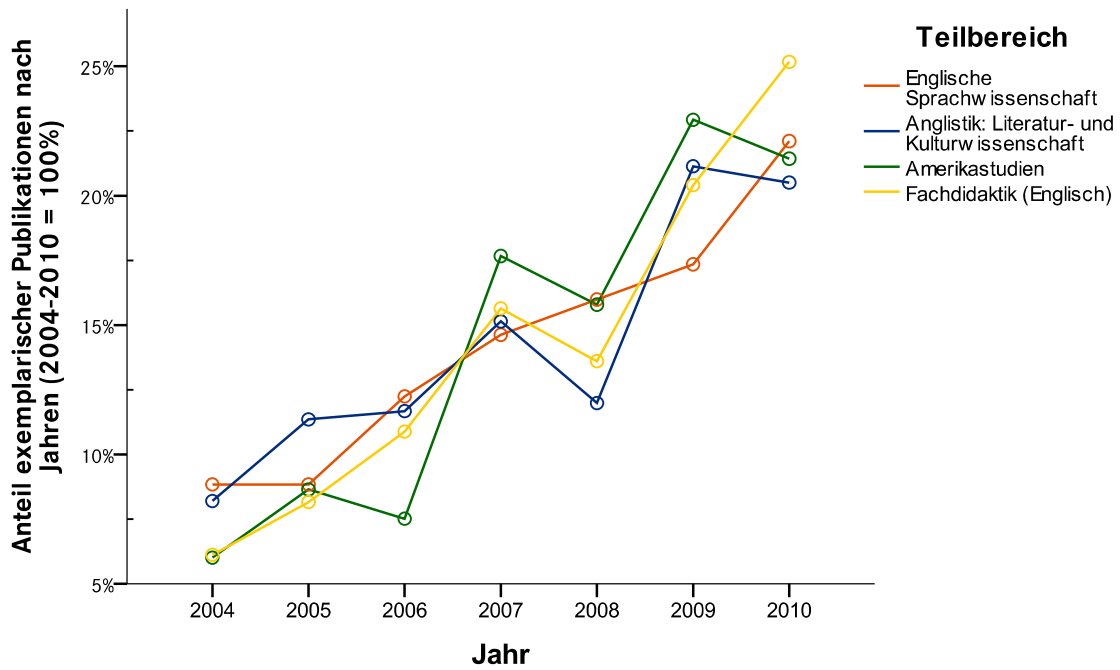
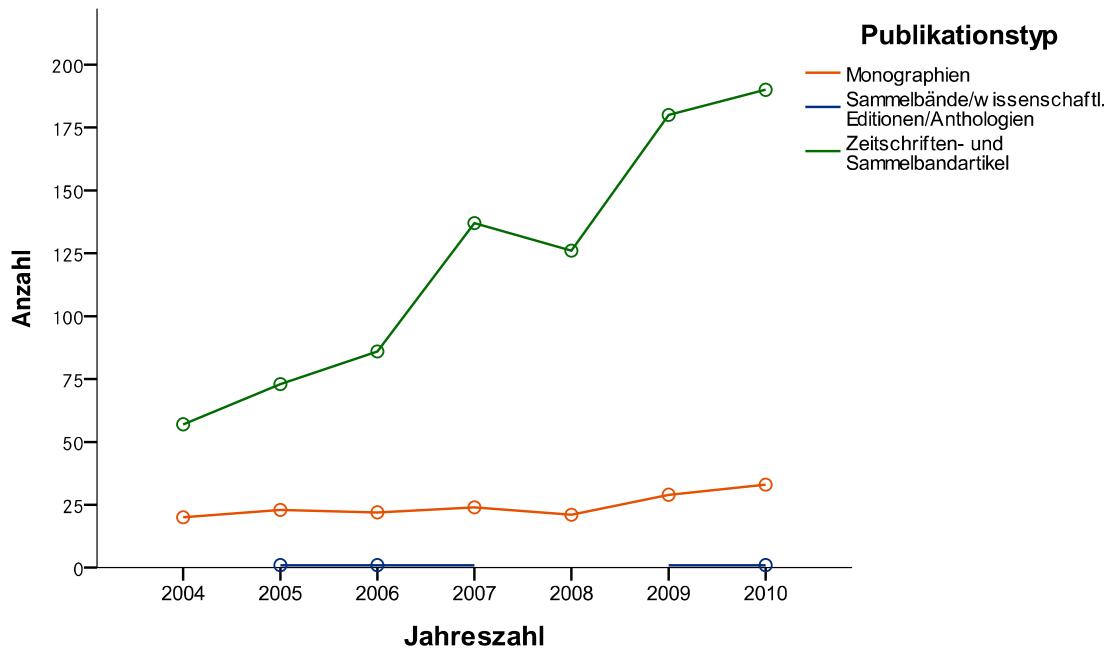


Abbildung 3: Anzahl exemplarischer Publikationen nach Publikationstyp und Jahren

Hauptdatenerhebung

- Die Hauptdatenerhebung fand von Mitte Oktober bis Ende Dezember 2011 statt.
- 5 Auch im Rahmen dieser Erhebung erreichten die Geschäftsstelle einige Nachfragen zur Bearbeitung der Fragebögen. Nach Abschluss der Datenerhebung wurden die Angaben der Einrichtungen auf Plausibilität und Konsistenz geprüft. An einigen Stellen sind dabei (scheinbare) Unstimmigkeiten aufgefallen bzw. haben sich Rückfragen ergeben. Soweit sich diese leicht aufklären ließen,
- 10 hat die Geschäftsstelle in Abstimmung mit den Einrichtungen die notwendigen Korrekturen vorgenommen. In denjenigen Fällen, in denen eine Klärung durch die Geschäftsstelle nicht möglich war, wurden die Einrichtungen um Korrekturen, Ergänzungen oder Präzisierungen gebeten. Sämtliche vorgenommenen Änderungen wurden dokumentiert und den Einrichtungen schriftlich mitgeteilt.
- 15 Die Benennung einer geeigneten Fachkoordinatorin bzw. eines geeigneten Fachkoordinators und die Unterstützung durch die Verwaltung spielten für den internen Ablauf der Datenerhebung eine wichtige Rolle. Auch das webgestützte Erhebungsinstrument hat sich bewährt und die im Rahmen der Pilotstudie Forschungsrating Elektrotechnik und Informationstechnik aufgetretenen techni-
- 20 schen Probleme konnten hier weitgehend behoben werden. |²¹

|²¹ Im Forschungsrating Elektrotechnik und Informationstechnik ergaben sich in einer Reihe von Fällen bedingt durch Zeilenbegrenzungen oder ungültige Steuerzeichen Probleme bei der Eingabe von Publikationslisten. So führten Formatierungsbesonderheiten (Absatzformate, Steuerzeichen etc. in den Listen dazu,

Die Geschäftsstelle hat alle erhobenen Daten auf Plausibilität und Konsistenz geprüft (siehe II.2) Die häufigsten Gründe für Rückfragen an die teilnehmenden Einrichtungen waren auch im Rahmen dieser Studie unvollständige oder fehlende Angaben oder Nennungen. Erst nach Abschluss der Datenprüfung und den erforderlichen Nacherhebungen bereitete die Geschäftsstelle für jede teilnehmende Einrichtung die vorliegenden Daten in einem so genannten Datenbericht auf, der sich aus zwei Teilen zusammensetzte: dem eigentlichen Datenbericht und den Publikationslisten. Die bei den Einrichtungen erhobenen Daten wurden partiell zu abgeleiteten Indikatoren verrechnet. Zusätzlich enthielten die Datenberichte die Rahmeninformationen der Einrichtung und ihrer Teilbereiche. Die Länge der Datenberichte hing hauptsächlich von der Anzahl der für die einzelnen Einrichtungen gemeldeten Teilbereiche ab und betrug im Mittel ca. 40 Seiten; das Minimum lag bei 15, das Maximum bei 70 Seiten. |²² Nach Eingang und Kontrolle aller Daten wurden für die quantitativen Indikatoren statistische Lagemaße (Perzentile) ermittelt und in die Datenberichte eingefügt. Weitere Verteilungsmaße (z. B. Median) wurden in einem „Leitfaden zu den Datenberichten für die Bewertungsgruppe Anglistik / Amerikanistik“ ausgewiesen, so dass sie in der Bewertung berücksichtigt werden konnten. Die fertig gestellten Datenberichte wurden auch den Einrichtungen (zunächst ohne Perzentilangaben) zur Kenntnis übersendet.

Die gewählte Form der Datenaufbereitung hat sich grundsätzlich bewährt und die Datenberichte boten einen übersichtlichen und klar strukturierten Datenzugang und eine gute Bewertungsgrundlage. Jedoch wurde auch im Rahmen dieser Studie deutlich, dass eine gründliche Datenkontrolle unverzichtbar ist und ausreichend Zeit benötigt. Trotz sorgfältiger Prüfung der eingegangenen Daten durch die Geschäftsstelle konnten nicht alle aus fachlicher Sicht unplausiblen oder unstimmgigen Angaben korrigiert werden. Wünschenswert wäre eine zusätzliche Kontrollsichtung der Daten durch die Gutachterinnen und Gutachter und eine entsprechende Rückfragerunde an die Einrichtungen, um die fehlerhaften oder lückenhaften Daten zu ergänzen.

Nach Einschätzung der Bewertungsgruppe Anglistik und Amerikanistik hat sich die Umrechnung der quantitativen Daten in skaleninvariante Lagemaße (Perzentile) grundsätzlich bewährt. Allerdings ist darauf hinzuweisen, dass solche Werte im Bewertungsprozess eine Komplexitätsreduktion darstellen, die zwin-

dass die Angaben zu einer Publikation auf mehrere Zeilen verteilt wurden, wobei jede Zeile als eine gesonderte Publikation gezählt wurde.

|²² Die Publikationslisten waren durchschnittlich 24 Seiten lang (min. 5 Seiten; max. 62 Seiten).

gend zu reflektieren ist. Wenn z. B. die im Leitfaden enthaltenen Maße zur Beschreibung der zugrunde liegenden Verteilung der Daten wie die Streuung nicht mit berücksichtigt werden, kann dies zu einer Überschätzung der Perzentildifferenzen führen. Daher bekräftigt auch die Bewertungsgruppe Anglistik und Amerikanistik die Vorbehalte im Forschungsrating hinsichtlich der Berechnung hoch aggregierter Zahlen, etwa gewichteter Gesamtindikatoren für einzelne Kriterien, da sie das vorhandene und für die Bewertung relevante Differenzierungspotenzial der einzelnen Daten zu weit reduzieren.

A.III BEWERTUNG

10 III.1 Definition der Bewertungsskala

Dem Beispiel der Bewertungsgruppe Elektrotechnik und Informationstechnik folgend, wurden die Bewertungen auf einer insgesamt neunstufigen Skala vorgenommen, die neben fünf ganzen Bewertungsstufen mit den Werten 5 = „herausragend“, 4 = „sehr gut“, 3 = „gut“, 2 = „befriedigend“ und 1 = „nicht befriedigend“ auch die jeweiligen Zwischenstufen |²³ umfasste. Auf diese Weise lassen sich differenzierte Bewertungsaussagen treffen.

Erfasst wurden die Bewertungen anhand webgestützter Bewertungsblätter, die analog zu der abgestimmten Bewertungsmatrix aufgebaut waren. Die Online-Bewertungsblätter wurden, wie auch der Online-Fragebogen, vom UZ entwickelt.

Bei der Veröffentlichung der Bewertungsergebnisse wurden die ganzen Bewertungsstufen mit den nachfolgenden Zwischenstufen farblich zusammengefasst (z. B. 5 und 4,5; 4 und 3,5 etc.).

III.2 Verfahren der Bewertung

25 Die Bewertung im Forschungsrating Anglistik und Amerikanistik wurde von einer Gruppe von 19 Fachvertreterinnen und Fachvertretern vorgenommen, von denen drei aus dem Ausland stammen. Darüber hinaus gehörte der Bewertungsgruppe ein fachfernes Mitglied der Steuerungsgruppe mit Gaststatus an. Die Bewertungsgruppe wurde vor Beginn des Bewertungsprozesses ergänzt und

|²³ 5-4 = „herausragend/sehr gut“, 4-3 = „sehr gut/gut“, 3-2 = „gut/befriedigend“, 2-1 = „befriedigend/nicht befriedigend“.

deckte daher im Rahmen des Bewertungsprozesses alle Bereiche ihrer Disziplin ab. |²⁴

Die Bewertungen wurden arbeitsteilig vorgenommen: Unter Berücksichtigung der Befangenheiten |²⁵ waren jeweils zwei Berichterstatterinnen und Berichter-
 5 statter für jede Einrichtung bzw. jeden Teilbereich zuständig. |²⁶ Sofern die eingereichten Publikationen inhaltlich in den benachbarten Teilbereich fielen, wurden diese auch Gutachterinnen und Gutachtern des entsprechenden Teilbereichs vorgelegt.

Ihre Ergebnisse legte die Bewertungsgruppe Anglistik und Amerikanistik in
 10 konsensorientierten Beratungen auf einer dreitägigen und zwei zweitägigen Sitzungen fest. Die Bewertung wurde in drei Stufen durchgeführt. Zunächst ermittelten die jeweils zuständigen Berichterstatterinnen und Berichterstatter unabhängig voneinander vorläufige Bewertungsvoten. Nachfolgend berieten die Gutachterinnen und Gutachter ihre vorläufigen Bewertungen in zwei Gruppen,
 15 die sich aus Gutachterinnen und Gutachtern von jeweils zwei Teilbereichen zusammensetzten (Anglistik: Literatur- und Kulturwissenschaft zusammen mit Amerikastudien sowie Englische Sprachwissenschaft zusammen mit Fachdidaktik Englisch). Für die Beratungen in den Teilbereichen hatte das UZ ein zusätzliches Online-Modul bereitgestellt, das jeweils die vorliegenden Bewertungen der
 20 beiden Gutachterinnen und Gutachter eines Teilbereichs gegenüberstellte. Dadurch wurden Übereinstimmungen und Abweichungen auf einen Blick sichtbar. Im Rahmen dieser Beratungen wurden alle Gutachternvoten behandelt und zu einem gemeinsamen Urteil zusammengeführt. Auf dieser Grundlage wurden alle Bewertungen zur Abstimmung im Plenum gestellt. Gesondert beraten wurde in den folgenden Fällen: Eine der beiden Extrembewertungen („herausragend“ bzw. „herausragend/sehr gut“ oder „befriedigend/nicht befriedigend“ bzw. „nicht befriedigend“) war vergeben worden; der Teilbereich bzw. die Einrichtung war nicht bewertet worden; ein oder mehrere Mitglieder der Bewertungsgruppe baten um Aussprache.

30 Im Unterschied zu den vorangegangenen Pilotstudien in Chemie, Soziologie und Elektrotechnik und Informationstechnik wurden im Forschungsrating Anglistik und Amerikanistik alle Kriterien auf der Ebene der Teilbereiche bewertet. Da-

|²⁴ Es kam eine Gutachterin im Teilbereich Amerikastudien sowie eine Gutachterin und ein Gutachter im Teilbereich Fachdidaktik hinzu, die Kompetenzen in bis dahin noch nicht abgedeckten Forschungsrichtungen mitbrachten.

|²⁵ Zudem wurden den beteiligten Gutachterinnen und Gutachtern keine Einrichtungen aus dem jeweiligen Sitzland ihrer Arbeitseinrichtung zugewiesen.

|²⁶ Die Mitglieder der Bewertungsgruppe wurden nach ihren fachlichen Schwerpunkten einzelnen Einrichtungen und Teilbereichen zugeordnet.

mit sollte der Verfasstheit des Fachs Anglistik und Amerikanistik Rechnung getragen werden. Nach der ersten Sichtung der Daten und in Vorbereitung der Bewertungsphase fanden Treffen der Gutachterinnen und Gutachter der jeweiligen Teilbereiche statt. Diese Treffen dienten der Entwicklung von teilbereichsspezifischen Bewertungsmaßstäben, die auf Basis von Daten zu in der Regel zwei Beispielenrichtungen beraten wurden. Diese teilbereichsspezifische Beratung der Bewertungsmaßstäbe vor der Durchführung der individuellen Vorbewertungen hatte in zweifacher Sicht positive Effekte und ist auch für die zukünftige Forschungsratings zu empfehlen. Zum einen stellte die frühzeitige Auseinandersetzung mit dem Datenmaterial und den Bewertungsmaßstäben innerhalb der jeweiligen Teilbereiche eine gemeinsame Bewertungsgrundlage sicher. Zum anderen konnte so ein für die einzelnen Teilbereiche der Anglistik und Amerikanistik jeweils angemessener Zugang zu Bewertung geschaffen werden. Vor Beginn der teilbereichsspezifischen Beratungen der Bewertungen wurden die in den teilbereichsspezifischen Treffen zu Beginn des Bewertungsprozesses jeweils festgelegten Maßstäbe in die anderen Teilbereiche kommuniziert und plenar diskutiert. Im Falle einer Wiederholung des Forschungsratings in der Anglistik und Amerikanistik sollte ein besonderer Wert darauf gelegt werden, die Ergebnisse im Anschluss an die teilbereichsspezifischen Treffen zusammenzuführen und die Bewertungsmaßstäbe in der Bewertungsgruppe gemeinsam zu definieren. Hilfreich hierfür wäre eine exemplarische Bewertung aller Teilbereiche an ein bis zwei Einrichtungen im Plenum.

Auch die Beratung der einzelnen Fälle in den Teilbereichen mit anschließender Plenumsberatung hat sich bewährt und sollte auch künftig beibehalten werden. Eine derartige Organisation des Bewertungsprozesses ermöglichte eine intensive Auseinandersetzung mit den einzelnen Fällen.

Nach Abschluss der Beratungen der einzelnen Fälle wurden Bewertungen aller Einrichtungen in den teilbereichsspezifischen Gutachtergruppen nochmals in einer Gesamtschau betrachtet und in diesem Lichte einzelne Entscheidungen gegebenenfalls korrigiert.

Die Grundlage der Bewertung bildeten die exemplarischen Publikationen, die Datenberichte sowie die Publikationslisten. Den Gutachterinnen und Gutachtern lagen die exemplarischen Publikationen der Teilbereiche, für die sie zuständig waren sowie Datenberichte und Publikationslisten aller teilnehmenden Einrichtungen und Teilbereiche vor, also auch derjenigen, für die sie nicht als Berichterstatterinnen bzw. Berichterstatter zuständig waren. Somit wurde in der plenaren Abstimmung eine für alle Mitglieder der Bewertungsgruppe nahezu einheitliche Informationsbasis sichergestellt. Nach jeder Bewertungssitzung wurden nach Einrichtungen sortierte Übersichten über die zum diesem Zeitpunkt erzielten Bewertungsergebnisse erstellt. Dies erwies sich insbesondere in Vorbereitung auf die letzte Sitzung, die der Konsolidierung der Bewertungen

und der Abstimmung über die „Extrembewertungen“ diene, als ein hilfreiches Mittel zur Strukturierung der Beratungen und zum Auffinden der Fälle, in denen größerer Diskussionsbedarf bestand.

5 Fallweise ist ein Kommentar der Bewertungsgruppe zur Bewertung eines jeweiligen Teilbereichs oder einer jeweiligen Einrichtung angegeben. Dieser gibt Aufschluss über etwaige Besonderheiten der Bewertung. Er kann damit der besseren Interpretation einer Bewertung dienen. Auch sind die Gründe für die „Nicht-Bewertung“ oder „Nicht-Veröffentlichung“ im jeweiligen Kommentar dokumentiert. Vor einer Veröffentlichung der Bewertungsergebnisse für Teilbereiche mit weniger als drei Wissenschaftlerinnen und Wissenschaftlern wurde aus datenschutzrechtlichen Gründen abgesehen. Insgesamt wurden 17 Teilbereiche nach einem oder mehreren Kriterien nicht bewertet. Diese Einschätzung enthält per se keine Wertung und wurde aus unterschiedlichen Gründen getroffen, beispielsweise wenn die Datengrundlage wenig aussagekräftig oder lückenhaft war. Weitergehende Analysen, insbesondere Vermutungen darüber, was die Leistungsfähigkeit der bewerteten Einrichtungen erklärt, können ohne eine tiefergehende Einzelevaluation nicht belegt werden. Eine systematische inhaltliche Kommentierung aller Teilbereiche und Einrichtungen wäre aufgrund der hohen zeitlichen Inanspruchnahme der Bewertungsgruppe durch die Beratungen im Bewertungsprozess nicht realisierbar.

Wie auch in den ersten drei Pilotstudien trat im Bewertungsverfahren eine unterschiedlich hohe Belastbarkeit der einzelnen Bewertungskriterien zutage. Die Daten zu den Kriterien „Forschungsqualität“ und „Forschungsermöglichung“ lieferten nach Einschätzung der Bewertungsgruppe sowohl in qualitativer als auch in quantitativer Hinsicht belastbare Bewertungsgrundlage. Die zur Bewertung der Kriterien „Reputation“ und „Transfer an außeruniversitäre Adressaten“ vorliegenden Daten waren hingegen häufig sehr heterogen oder lückenhaft.

Auch die Qualität der Forschungsleistungen der einzelnen Wissenschaftlerinnen und Wissenschaftler innerhalb eines Teilbereichs einer Einrichtung war zum Teil sehr heterogen. Für die Weiterentwicklung des Verfahrens ist die Frage bedeutsam, wie man mit der Leistungsvarianz innerhalb der Teilbereiche umgeht. Besonders wichtig wäre die Analyse der Heterogenität bei Teilbereichen mit mittlerer Leistungsstärke, um fundierte Anhaltspunkte zu Bedingungsfaktoren guter Forschungsqualität zu erhalten. In diesem Zusammenhang sollte die Frage der Granularität der Forschungseinheiten gezielt mit den Adressaten des Forschungsratings diskutiert werden.

Das Verfahren der Bewertung hat sich grundsätzlich bewährt und sollte bei einer Wiederholung beibehalten werden. Das zur Anwendung gebrachte *informed-peer-review* zeigt sich als robust und liefert verlässliche Ergebnisse. Die von jeweils zwei Gutachterinnen und Gutachtern auf Grundlage der diskutierten Be-

wertungsmaßstäbe, jedoch unabhängig voneinander vergebenen Vorbewertungen der insgesamt neun Bewertungsaspekte und vier Bewertungskriterien, korrelieren hoch ($\rho = 0,8$, $p < 0,01$, $N = 4110$). 40 Prozent der Bewertungen von jeweils zwei Gutachterinnen und Gutachtern waren identisch, bei weiteren 40 Prozent der Bewertungen bestand lediglich eine Differenz von 0,5 (Mittelwert der Differenzen = 0,46; St.abw. = 0,5; Median = 0,5). Die höchste Übereinstimmung der Bewertungen wurde beim Bewertungsaspekt „Drittmittelaktivitäten“ (Kriterium Forschungsermöglichung) erreicht. Die Gesamtbewertungen des Kriteriums Forschungsqualität, die überwiegend auf der qualitativen Beurteilung der eingereichten Publikationen und Publikationsauszüge fußten, weisen die zweithöchste Übereinstimmung der Bewertungen zwischen den beiden Gutachterinnen und Gutachtern auf. Dies zeigt, dass in der Anglistik und Amerikanistik relativ einheitliche Standards für die Beurteilung von Publikationen existieren. Als inhomogener erwiesen sich hingegen die Bewertungen des Aspekts „Wissensvermittlung“ (Kriterium Transfer an außeruniversitären Adressaten). Ungeachtet der hohen Korrelationen sieht die Bewertungsgruppe keine nennenswerten Möglichkeiten, den Aufwand der Datenerhebung und -bewertung zu verringern.

III.3 Bewertungsmatrix

In der Bewertungsmatrix wurde das Ergebnis der Operationalisierung zusammengefasst und damit die Struktur sowohl für die Datenerhebung als auch für die Bewertungsphase festgelegt. Der fachspezifischen Ausgestaltung einer solchen Matrix kommt daher besonderes Gewicht zu.

A.IV VERÖFFENTLICHUNG DER ERGEBNISSE

Nach Abschluss des Bewertungsprozesses wird die Bewertungsgruppe die Ergebnisse des Forschungsratings Anglistik und Amerikanistik der für das Verfahren verantwortlichen Steuerungsgruppe vorlegen. Vorab werden den Fachkoordinatorinnen und Fachkoordinatoren die Ergebnisse ihrer Einrichtung zur Kenntnis zugesandt. Damit wird dem Wunsch der Fachgemeinschaft entsprochen, kommentierende Hinweise abzugeben, die aus der Sicht der Einrichtungen für die Beratungen der Steuerungsgruppe relevant sein könnten.

Vorbehaltlich der Verabschiedung durch die verantwortliche Steuerungsgruppe werden die Ergebnisse des Forschungsratings Anglistik und Amerikanistik am 7. Dezember 2012 durch die Steuerungsgruppe veröffentlicht. Wie in den vorangegangenen Pilotstudien erfolgt die Veröffentlichung im Rahmen einer Pressekonzferenz. Zum selben Zeitpunkt wird auch der Zugang zu Ergebnissen auf der Homepage des Wissenschaftsrats (www.forschungsrating.de) eröffnet. Im Rahmen der Ergebnisveröffentlichung wird zusätzlich der Ergebnisbericht der

Bewertungsgruppe präsentiert. In diesem werden die inhaltlichen und organisatorischen Züge des Verfahrens dargestellt. Darüber hinaus wird die Ergebnisdarstellung für die teilnehmenden Einrichtungen vorgestellt. Für jede Einrichtung wird hierzu ein individuelles Datenblatt erstellt, das eine teilbereichsbezogene und nach Bewertungskriterien differenzierte Ergebnisdarstellung enthält. Zudem werden folgende einrichtungsbezogenen Informationen dargestellt:

5

Einrichtungstyp: Es wird angegeben, ob es sich um eine Universität oder eine Pädagogische Hochschule handelt.

10 Internetadresse(n)

Rahmeninformationen:

Unterteilt in jeweils vier Größenklassen wird angegeben:

_ Kopfzahl des gesamten wissenschaftlichen Personals |²⁷ (< 3; 3 – 5; 6 – 9; > 9),

_ davon Professuren (≤ 1; 2 – 3; > 3)

15 _ Studienfälle im Hauptfach (ohne Lehramt) (< 250; 250 – 400; 401 – 700; > 700)

_ Studienfälle Lehramt (< 400; 400 – 700; 701 – 1300; > 1300)

_ Abschlussprüfungen im Haupt- und Nebenfach 2004 – 2010 |²⁸ (< 550; 550 – 900; 901 – 1400; > 1400; k. A.);

20 Vakanzen zum Stichtag 31.12.2010: Die von den Einrichtungen im Rahmen der Datenerhebung angegebenen Vakanzen.

25 Kommentar zur Bewertung: Fallweise wird ein Kommentar der Bewertungsgruppe zur Bewertung einer jeweiligen Einrichtung bzw. eines Teilbereichs angegeben. Dieser gibt Aufschluss über etwaige Besonderheiten der Bewertung. Darüber hinaus wurden die Gründe für die „Nicht-Bewertung“ oder „Nicht Veröffentlichung“ dokumentiert.

Bewertungsergebnisse der einzelnen Teilbereiche in Matrixform:

|²⁷ Professorinnen und Professoren, habilitierte Wissenschaftlerinnen und Wissenschaftler, Wissenschaftliche Mitarbeiterinnen und Mitarbeiter. Nicht ausgewiesen ist Personal in der Lehre (Lehrkräfte für besondere Aufgaben, Lektoren, Mitarbeiter in der Sprachpraxis) sowie nicht wissenschaftliches Personal (z. B. Sekretariate).

|²⁸ Staatsexamen, Magister, Bachelor, Master

Bewertungs- kriterium Teilbereich	I. Foschungsqua- lität	II. Reputa- tion	III. Forschungs- ermöglichung	IV. Transfer an außeruniversi- täre Adressaten
Englische Sprach- wissenschaft				
Anglistik: Literatur- und Kulturwissen- schaft				
Amerikastudien				
Fachdidaktik Eng- lisch				

Die Ergebnispräsentation auf der Homepage des Wissenschaftsrates bietet neben einrichtungsbezogenen Datenblättern einen Vergleich der Teilbereiche für die einzelnen Bewertungskriterien.

Bewertungsergebnisse für Teilbereiche, in denen weniger als 3 Wissenschaftlerinnen und Wissenschaftler gemeldet wurde sowie für Teilbereiche, in denen keine Professur gemeldet wurde, werden aus datenschutzrechtlichen Gründen nicht veröffentlicht. Die Einrichtungen selbst haben jedoch die Möglichkeit, diese Ergebnisse mit der Einwilligung der Betroffenen zu veröffentlichen. Die Einrichtungen und die Länder erhalten ihre eigenen Bewertungsergebnisse eine Woche vor der Veröffentlichung. Die Einrichtungen selbst erhalten außerdem den jeweiligen „Datenbericht“, der die Bewertungsgrundlage bildete. Dieser erfasst die in den Fragebögen und aus externen Quellen erhobenen Daten zu den einzelnen Kriterien und gibt für die quantitativen Daten Verteilungskennwerte an. Die Datenberichte sind vertraulich und werden daher nicht veröffentlicht, können jedoch von den Einrichtungen für ihre Zwecke genutzt werden.

Es ist geplant, die Ergebnisse des Forschungsratings auch auf Tagungen der Fachgesellschaften im Jahr 2013 vorzustellen.

B. Zusammenfassende Bewertung und Folgerungen

B.1 ZUSAMMENFASSENDE BEWERTUNG VON AUFWAND UND ERTRAG

5 Bei einer generellen Einschätzung von Aufwand und Ertrag ist der Pilotcharakter der Studie zu beachten. Die Pilotphase hat eine Reihe von zu erprobenden Komponenten und Absicherungsmaßnahmen erfordert, die bei einer Verstetigung entfallen und folglich den Aufwand reduzieren können. Die wesentlichen
10 Träger des Aufwands sind die Mitglieder der Bewertungsgruppe, die teilnehmenden Einrichtungen und die Geschäftsstelle. Nachfolgend werden diejenigen Faktoren skizziert, die den Aufwand an den jeweiligen Stellen beeinflussen.

Bewertungsgruppe

Die 19 Mitglieder |²⁹ der Bewertungsgruppe Anglistik und Amerikanistik sind während der Pilotstudie im Zeitraum März 2011 bis Oktober 2012 zu insgesamt
15 acht Sitzungen (13 Sitzungstage) |³⁰ zusammengetroffen. Drei davon dienten der fachspezifischen Anpassung des Verfahrens, vier Sitzungen der Bewertungsphase einschließlich Vorbereitung. Die letzte Sitzung diente der Abstimmung des vorliegenden Abschlussberichts. Zusätzlich zu den regulären Sitzungen hat ein Teil der Bewertungsgruppe sich für einen Tag zur abschließenden
20 Redaktion der Fragebögen getroffen. Darüber hinaus haben pro Teilbereich ein-

|²⁹ Zudem gehörte der Bewertungsgruppe neben den 19 Gutachterinnen und Gutachern ein fachfernes Mitglied der Steuerungsgruppe mit Gaststatus an.

|³⁰ Im Forschungsrating Chemie traf sich die Bewertungsgruppe zu insgesamt acht Sitzungen (elf Sitzungstage), im Forschungsrating Soziologie zu elf Sitzungen (15 Sitzungstage), im Forschungsrating Elektrotechnik und Informationstechnik zu elf Sitzungen (15 Sitzungstage).

tägige Treffen zur Entwicklung der Bewertungsmaßstäbe stattgefunden. Hinzu kommt die individuelle Vorbereitung, die vor allem in der Bewertungsphase einen erhöhten Aufwand bedeutete. Die Gutachterinnen und Gutachter haben insgesamt zwischen drei und vier Monate für die Durchführung des Forschungsratings in der Anglistik und Amerikanistik aufgewendet, von denen bis zu zwei Monate auf die Lektüre der exemplarischen Publikationen entfielen.

Der hohe Aufwand der Gutachterinnen und Gutachter kann teilweise auf die Entwicklung und Erprobung des Verfahrens zurückgeführt werden. So mussten zunächst grundlegende Operationalisierungsfragen geklärt, die Erhebungsinstrumente entwickelt und die Bewertungsskala festgelegt werden. Dies erfolgte größtenteils im Zeitraum März bis Juni 2011. Ausgehend von den vorliegenden Erfahrungen der Pilotstudie könnte bei einer Wiederholung des Verfahrens der Aufwand für die fachspezifische Operationalisierung durch ein höheres Standardisierungsniveau reduziert werden. Ganz verzichtbar ist die Operationalisierungsphase jedoch nicht, da die Kriterien und festgelegten Indikatoren bei einem mehrjährigen Turnus auch dann überprüft und gegebenenfalls aktualisiert werden müssten, wenn dasselbe Fach erneut bewertet würde. Zudem beförderte der Prozess der Operationalisierung auch die Auseinandersetzung mit der komplexen Methodik in der Bewertungsgruppe.

Der hohe Aufwand der Gutachterinnen und Gutachter kann darüber hinaus auf die zeitintensive Lektüre der exemplarischen Publikationen zurückgeführt werden. Jeder Gutachterin und jedem Gutachter lagen im Mittel 110 Publikationen (min. 67, max. 133) vor.³¹ Die Durchschnittliche Publikationslänge betrug knapp 24 Seiten (min. 2 Seiten, max. 50 Seiten). Insgesamt hing der Arbeitsaufwand in der Bewertungsphase von der Anzahl der zu bewertenden Teilbereiche, der Anzahl der Bewertungskriterien und der Qualität der Daten ab. In der Pilotstudie Anglistik und Amerikanistik hatte jedes Mitglied der Bewertungsgruppe als Berichterstatterin bzw. als Berichterstatter die Verantwortung für im Mittel knapp 21 Teilbereiche, die jeweils nach vier Kriterien zu bewerten waren. Jedes Mitglied musste also insgesamt über 84 Einzelbewertungen abgeben. Betrachtet man die Ebene der Bewertungsaspekte, dann wird die Dimension des Bewertungsaufwandes mit im Mittel 189 einzelnen Bewertungen noch deutlicher. Ne-

³¹ Die Anzahl der zu lesenden Publikationen variierte stark und war abhängig von der Anzahl der gemeldeten Teilbereiche und Professuren: Englische Sprachwissenschaft: min. 118, max. 133 Publikationen; Anglistik: Literatur- und Kulturwissenschaft: min. 119, max. 130 Publikationen; Amerikastudien: min. 92; max. 122 Publikationen; Fachdidaktik Englisch: min. 67, max. 83 Publikationen). Bei der Zuteilung der zu bewertenden Teilbereiche und entsprechend der Publikationen wurde neben der fachlichen Kompetenz und Befangenheiten auch der Umfang der zu lesenden Publikationen berücksichtigt, so dass die Gutachterinnen und Gutachter einzelner Teilbereiche unabhängig von der Anzahl der Publikationen einen in etwa gleich großen Lektüreaufwand hatten.

ben inhaltlichen Gründen spricht auch dies dafür, den Differenzierungsgrad der Bewertung zu begrenzen. Die methodische Reflexion des Verfahrens im vorliegenden Abschlussbericht ist ebenfalls spezifisch für eine Pilotstudie.

Bewertete Einrichtungen / Datenerhebung

5 Das *informed-peer-review*-Verfahren setzt eine aussagekräftige Datengrundlage voraus, die größtenteils bei den teilnehmenden Einrichtungen zu erheben ist. Der Arbeitsaufwand hierfür lag vor allem bei den zuständigen Fachkoordinatorinnen und Fachkoordinatoren. Neben der Größe einer Einrichtung zählen auch die Qualität des internen Controllings und die Unterstützung der Wissenschaftlerinnen und Wissenschaftler durch die Verwaltung zu den Faktoren, die den
10 Arbeitsaufwand wesentlich beeinflussen.

Ca. fünf Wochen vor Beginn der Datenerhebung wurden zwei Informationsveranstaltungen für die Fachkoordinatorinnen und Fachkoordinatoren angeboten. Bei beiden Veranstaltungen haben die Teilnehmenden ausgiebig von der Möglichkeit Gebrauch gemacht, Fragen zu stellen und Änderungsvorschläge zu unterbreiten, die, soweit möglich, in die Fertigstellung der Fragebögen eingingen. Diese frühzeitige Information der Einrichtungen zum Erhebungsverfahren hat sich bewährt; die so erreichte längere Vorlaufzeit und auch die komfortable Handhabung des Online-Erhebungsinstruments haben erheblich zu einer Optimierung des Verfahrens beigetragen.
15
20

Ein weiterer zentraler Faktor ist der Inhalt der Datenerhebung. Einige Daten (z. B. Daten zu Bewertungsaspekten „Anerkennung“ und „Personaltransfer“, ausführlicher hierzu vgl. Kapitel A.I.2 dieses Berichts) haben sich in der Pilotstudie als wenig belastbar herausgestellt und sollten nicht wieder erhoben werden.

25 Darüber hinaus spielte eine Rolle, dass nur wenige Daten auf der Ebene der Einrichtung, der Großteil jedoch nach Teilbereichen differenziert erhoben wurden. Relevante Angaben auf dieser Aggregationsstufe, die partiell erst für das Forschungsrating definiert bzw. abgegrenzt wurden, lagen nicht bei allen Einrichtungen vor und mussten in diesen Fällen aus Angaben der teilnehmenden Wissenschaftlerinnen und Wissenschaftler neu zusammengestellt werden. Dieses Problem ist dann zu lösen, wenn die Einrichtungen kontinuierlich die wichtigsten Daten zentral vorhalten. Der Wissenschaftsrat erarbeitet derzeit hierzu Empfehlungen. Diese werden sowohl einen Vorschlag eines Datensets, das von den meisten Akteuren verwendet und von den Hochschulen und außeruniversitären Forschungseinrichtungen gepflegt und vorgehalten werden soll, als
30
35 auch Empfehlungen zum Implementierungsprozess beinhalten.

Die Pilotstudie Forschungsrating Anglistik und Amerikanistik wurde in der Geschäftsstelle des Wissenschaftsrates von einer Projektgruppe mit insgesamt fünf Mitarbeiterinnen und Mitarbeitern unterstützt, die parallel die Pilotstudie Elektrotechnik und Informationstechnik, die Unterarbeitsgruppe „Vergleichende Forschungsbewertung in den Geisteswissenschaften“ sowie die Unterarbeitsgruppe „Standardisierung der Datenerhebung“ und die Arbeitsgruppe „Kerndatensatz Forschung“ betreute. Neben der Administration des gesamten Verfahrens wirkte die Projektgruppe an der Entwicklung von Indikatoren und der Konzeption der Erhebungsinstrumente mit. Die Datenerhebung erfolgte in enger Abstimmung mit der Geschäftsstelle, bei der auch die Kontrolle und Aufbereitung der Daten lag. Die Geschäftsstelle übernahm darüber hinaus die Prüfung und Bibliographie der eingereichten exemplarischen Publikationen. Eine ausreichende Kapazität der Geschäftsstelle ist für die Qualität der Verfahrensentwicklung und -organisation sowie der Datenerhebung essentiell und trägt nicht zuletzt zur Entlastung der Bewertungsgruppe bei.

B.II WEITERES VORGEHEN

Die Ergebnisse der Pilotstudie Anglistik und Amerikanistik zeigen, dass eine dem Fach angemessene vergleichende Bewertung von Forschungsleistungen der Anglistik und Amerikanistik grundsätzlich möglich ist. Das Forschungsrating bietet einen geeigneten Rahmen für die vergleichende Bewertung von Forschungsleistungen geisteswissenschaftlicher Fächer. Fachliche Besonderheiten wie z. B. die Verfasstheit des Fachs oder Publikationsaktivitäten, konnten durch die Ausgestaltung des Bewertungsmodells und der resultierenden Bewertungsmatrix sowie durch Organisation des Bewertungsprozesses angemessen berücksichtigt werden. Die hohe Übereinstimmung der Bewertungen zwischen den Gutachterinnen und Gutachtern zeugen von der Verlässlichkeit des Verfahrens und der Methode des *informed-peer-review*. Insbesondere hat sich die Bewertung auf der Ebene der Teilbereiche bewährt. Eine Einbeziehung von einigen wenigen für das Fach relevanten außeruniversitären Forschungseinrichtungen wäre wünschenswert gewesen. Grundsätzlich wird eine Wiederholung des Forschungsratings in der Anglistik und Amerikanistik unter der Berücksichtigung der in diesem Bericht angesprochenen Anpassungen empfohlen.

DFG	Deutsche Forschungsgemeinschaft
FhG	Fraunhofer-Gesellschaft
HGF	Helmholtz-Gemeinschaft Deutscher Forschungszentren
HRK	Hochschulrektorenkonferenz
MPG	Max-Planck-Gesellschaft
UZ	Umfragezentrum Bonn - Prof. Rudinger GmbH - Gesellschaft für empirische Sozialforschung und Evaluation
WGL	Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz
WR	Wissenschaftsrat

STATISTISCHES BUNDESAMT: Bildung und Kultur, Personal an Hochschulen, Fachserie 11, Reihe 4.4, Wiesbaden 2010.

WISSENSCHAFTSRAT: „Empfehlungen zu Rankings im Wissenschaftssystem Teil 1: Forschung“ (Drs. 6285-04), Hamburg 2004.

WISSENSCHAFTSRAT: Empfehlungen zum Forschungsrating (Drs. 8485-08), Rostock 2008.

WISSENSCHAFTSRAT: Pilotstudie Forschungsrating. Empfehlungen und Dokumentation, Köln 2008.

WISSENSCHAFTSRAT: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften, Köln 2010.

WISSENSCHAFTSRAT: Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011.

WISSENSCHAFTSRAT: Ergebnisse des Forschungsratings Anglistik und Amerikanistik (Drs. 2756-12), Köln 2012.

Abbildung 1: Anzahl exemplarischer Publikationen nach Teilbereichen und Jahren	35
Abbildung 2: Anteil exemplarischer Publikationen nach Teilbereichen und Jahren	35
Abbildung 3: Anzahl exemplarischer Publikationen nach Publikationstyp und Jahren	36