

Drs. 1656-11
Halle 11 11 2011

Empfehlungen zur Bewertung und Steuerung von Forschungsleistung

Vorbemerkung	5
A. Einleitung	7
A.I Anlass und Gegenstand	7
A.II Positionsbestimmung	10
A.III Ziele der Empfehlungen	13
B. Beobachtungen	15
I.1 Zur Bewertung durch <i>peers</i>	16
I.2 Zur Bewertung mit Hilfe quantitativer Indikatoren	19
I.3 Zu Evaluationsverfahren	21
I.4 Zu Ratings/Rankings	23
I.5 Zur Steuerung über Mittelallokation	24
I.6 Zu den Standards guter wissenschaftlicher Praxis	30
I.7 Zu den Folgen für die Lehre	31
I.8 Zu den wissenschaftspolitischen Rahmenbedingungen	32
C. Empfehlungen	33
C.I Übergeordnete Empfehlungen	33
C.II Einzelne Empfehlungen	35
II.1 Zum Einsatz komplementärer Steuerungszugänge	35
II.2 Zur Verwendung von <i>peer review</i>	37
II.3 Zu Indikatoren für Forschungsleistung	38
II.4 Zu Evaluationsverfahren	42
II.5 Zu Ratings/Rankings	44
II.6 Zu Verfahren der Mittelallokation	44
II.7 Zum Dokumentationswesen	48

Vorbemerkung

Der Wissenschaftsrat hat seit den 1980er Jahren mehrfach und in unterschiedlichen Kontexten zu Fragen der Leistungsdifferenzierung, des Wettbewerbs sowie zu Verfahren der Bewertung und Steuerung von Forschungsleistung Stellung genommen. Zugleich gestaltet er das Wissenschaftssystem durch Evaluations- und Akkreditierungsverfahren, das Forschungsrating und die Exzellenzinitiative an wesentlichen Stellen mit.

Inzwischen nimmt der Wissenschaftsrat wahr, dass viele Wissenschaftlerinnen und Wissenschaftler die Verfahren der Leistungsmessung und die Formen der Mittelallokation mit Unmut betrachten und hinsichtlich des Nutzens dieser Verfahren für die Wissenschaft zunehmend Skepsis äußern. Der Wissenschaftsrat erachtet es daher als notwendig, diese Debatten aufzugreifen, die Chancen und Risiken der jüngeren Entwicklungen herauszuarbeiten und Empfehlungen zu geben, die dem übergeordneten Ziel der Ermöglichung guter Forschung gerecht werden können.

Der Wissenschaftsrat ist sich bewusst, dass Veränderungen in der Forschungspraxis, die häufig auf Verfahren der Steuerung und Bewertung von Forschungsleistung zurückgeführt werden, vielfältige und vor allem auch systembedingte Ursachen haben können, wie etwa: Das Größenwachstum des Wissenschaftssystems, die wachsenden externen Ansprüche an die Wissenschaft, die Internationalität der Forschung, die Herausforderungen des Föderalismus, die komplexen Finanzierungswege, die zunehmende Angewiesenheit auf Drittmittel, die nicht immer gegebene Nachhaltigkeit wissenschaftspolitischer Initiativen, die Strukturierung des Systems in einen hochschulischen und einen außeruniversitären Sektor. Diese Voraussetzungen bilden den Hintergrund, vor dem der Wissenschaftsrat seine Empfehlungen abgibt.

Zur Vorbereitung der Empfehlungen hat der Wissenschaftsrat eine Arbeitsgruppe eingesetzt, die im September 2010 ihre Arbeit aufgenommen hat. In dieser Arbeitsgruppe haben auch Sachverständige mitgewirkt, die nicht Mitglieder des Wissenschaftsrates sind. Ihnen ist der Wissenschaftsrat zu besonderem Dank verpflichtet. Um ein möglichst differenziertes Bild der aktuellen Situation zu erhalten, hat die Arbeitsgruppe mit Vertreterinnen und Vertretern

6 der Wissenschaftsministerien ausgewählter Länder, großer Wissenschaftsorganisationen sowie des wissenschaftlichen Nachwuchses Gespräche zur Frage der wesentlichen Probleme und der Wirksamkeit der bestehenden Steuerungs- und Bewertungsverfahren geführt. Der Wissenschaftsrat dankt allen Beteiligten für ihre Unterstützung.

Der Wissenschaftsrat hat die Empfehlungen zur Bewertung und Steuerung von Forschungsleistung am 11. November 2011 in Halle an der Saale verabschiedet.

A. Einleitung

A.1 ANLASS UND GEGENSTAND

Leistungsmessung und -bewertung sowie die darauf aufbauenden Versuche der Steuerung von Forschung und deren Wirkungen sind seit einigen Jahren Gegenstand intensiver Debatten in Wissenschaft und Wissenschaftspolitik. Befürworter neuer Steuerungsansätze, die Hochschulen und außeruniversitäre Forschungseinrichtungen in ihrer strategischen Handlungsfähigkeit aufwerten wollen, stehen Verfechtern traditioneller akademischer Ordnungsprinzipien gegenüber, die das forschende Individuum als Nukleus der Forschung in Gefahr sehen. Die konträren Positionen lassen sich – zugespitzt und unter Ausklammerung von moderaten Zwischentönen – wie folgt beschreiben:

Die Befürworterinnen und Befürworter von neuen Steuerungsverfahren in der Forschung betonen, dass die Verteilung knapper öffentlicher Ressourcen einer nachvollziehbaren, transparenten Begründung gegenüber den politisch Verantwortlichen wie auch den betroffenen Wissenschaftlerinnen und Wissenschaftlern bedarf. Umfangreiche Investitionen in die Wissenschaft legitimieren sich demnach am ehesten durch herausragende und nützliche Forschungsergebnisse. Diese Position geht ferner davon aus, dass es messbare Leistungsunterschiede zwischen den Forschungseinrichtungen wie auch unter den Forschenden gibt. Sie fordert eine Steigerung von Leistung und Effizienz durch einen gezielten Ressourceneinsatz und befürwortet die weitere Differenzierung der Institutionenlandschaft. Um den Wettbewerb zu fördern, soll der Staat sich aus der Detailsteuerung der wissenschaftlichen Einrichtungen zurückziehen und stattdessen aus einer übergeordneten Perspektive und auf informierter Basis gezielte Impulse setzen. Die „Kosten“ für den Zugewinn an Autonomie auf Seiten der Wissenschaft bestehen in der Verpflichtung zu größerer Transparenz und Rechenschaftslegung. Strategisch handlungsfähig ist in dieser Position in erster Linie die autonome Institution. Zur Steigerung des Wettbewerbs soll ein Tableau von ziel- und leistungsbezogenen Maßnahmen eingesetzt werden, das Unterschiede in der Leistungsfähigkeit transparent macht, Anreize für ein bestimmtes Verhalten setzt – und die Akteure schließlich auch belohnt oder sanktioniert. Die Befürworter sehen ein motivationsförderndes Moment darin,

dass „gute“ wissenschaftliche Leistungen auch materiell belohnt werden können, wobei vorausgesetzt wird, dass eine gezielte institutionelle Ressourcenverteilung anhand valider Kriterien möglich ist. Diese Position hat sich in der Wissenschaftspolitik weitestgehend durchsetzen können und prägt heute die Rahmenbedingungen des Wissenschaftssystems. Wesentliche mit dieser Position verbundene Schlagworte sind: institutionelle Autonomie, Wettbewerb, leistungsorientierte Mittelverteilung, *New Public Management*, Effizienz, Evaluation, Transparenz.

Die Kritikerinnen und Kritiker hingegen sehen mit der Einführung von neuen Instrumenten der Bewertung und Steuerung das klassische Ideal der aus eigenem Antrieb selbstbestimmt Forschenden in Frage gestellt. Die einzelne Wissenschaftlerin bzw. der einzelne Wissenschaftler benötigt demnach grundsätzlich weder Überwachung noch Disziplinierung, sondern ist dann am leistungsstärksten, wenn sie oder er Vertrauen sowie ausreichend Freiräume und Ressourcen für ihre bzw. seine Forschungstätigkeiten erhält. Vertrauen, Zeitautonomie und intrinsische Motivation sind die Schlüssel zu leistungsstarker und kreativer Forschung. Wettbewerb im Wissenschaftssystem besteht primär im Wettbewerb um neue Forschungsergebnisse und deren Anerkennung in der *scientific community*. Zusätzliche externe, vor allem monetäre wettbewerbliche Impulse, die durchsetzungsstärkeren Forschenden ein höheres persönliches Einkommen oder eine bessere Ausstattung der Arbeitsumgebung versprechen, sind nicht erforderlich, um die Forschungsleistung zu steigern – im Gegenteil können sie sogar kontraproduktiv wirken und die intrinsische Motivation schwächen bzw. die entsprechenden sozialen Voraussetzungen dafür zerstören. Neben Effekten auf einzelne Wissenschaftlerinnen und Wissenschaftler werden auch negative Effekte für das Wissenschaftssystem als Ganzes befürchtet, die sich vielfach erst langfristig bemerkbar machen könnten, beispielsweise in der Vernachlässigung der Forschungsqualität zugunsten der Forschungsquantität, der Tendenz zu risikoarmer Forschung im Mainstream, der abnehmenden Diversität der Forschenden und Forschungsthemen sowie der zunehmenden Kleinteiligkeit der publizierten Forschungsergebnisse. Vorzuziehen ist dieser Position zufolge daher ein Zustand, in dem angemessene Rahmenbedingungen gesetzt werden, innerhalb derer sich die Forschenden autonom entfalten können. Der seitens der Geldgeber gestellte Anspruch an Rechenschaftslegung, Transparenz und Evaluation wird verrechnet mit unbeabsichtigten Nachteilen, die diese bergen: Steigerung der Konkurrenz bis zum Misstrauen, Demotivation der „Verlierer“, unangemessener Anstieg des Antrags- und Evaluationsaufwandes. Bemühungen der Wissenschaftspolitik und der Leitungen wissenschaftlicher Einrichtungen, steuernd in den wissenschaftlichen Prozess einzugreifen, werden deshalb als wissenschaftsfremder Anspruch abgelehnt. Wesentliche Schlagworte dieser

Position sind: individuelle Autonomie, Vertrauen, intrinsische Motivation, Forschungsfreiraum, Kreativität, Beharrlichkeit.

Diese zugespitzt dargestellte Auseinandersetzung um Bewertung und Steuerung der Forschungsleistung und ihre Effekte findet ebenso in Fachpublikationen und auf Tagungen wie in den Feuilletons statt. Die dabei geäußerten Positionen hängen mit grundlegenden Differenzen in der Bewertung und Wahrnehmung des heutigen Wissenschaftssystems zusammen, die sich mit den Schlagworten „*New Public Management*“ auf der einen und „Humboldt-Ideal“ auf der anderen Seite verbinden lassen. Beide Pole dieser Debatte gilt es zunächst in einen Kontext zu stellen, um anschließend die Zielspannungen herauszuarbeiten, die jede Art der Leistungsmessung und Forschungsbewertung hervorbringt.

Ansatzpunkt der skizzierten Debatten sind substantielle Veränderungen in den Rahmenbedingungen wissenschaftlichen Arbeitens. Dabei sind für das deutsche Wissenschaftssystem vor allem folgende Entwicklungen seit den 1980er Jahren prägend gewesen: auf internationaler Ebene die Einführung des *New Public Management* in der Wissenschaftspolitik der angelsächsischen Welt; auf nationaler Ebene der Eindruck eines Zurückfallens der Bundesrepublik Deutschland im internationalen Vergleich der Forschungsleistungen in vielen Disziplinen trotz des zwei Jahrzehnte andauernden Ausbaus des Wissenschaftssystems. Ferner nahm die Grundfinanzierung der Hochschulen nicht mehr im gleichen Maße zu wie die an sie gestellten Anforderungen. Als Kompensation für eine mangelnde Grundfinanzierung sowie als Leistungs- und Qualitätsindikator erhielten Drittmittel in allen Bereichen des deutschen Wissenschaftssystems ein wachsendes Gewicht. Schließlich ist die internationale und fächerübergreifende Ausbreitung indikatorengestützter Rankings und formalisierter Evaluationsverfahren zu nennen. Beide stehen im Zusammenhang mit einem gestiegenen Steuerungs- und Rechenschaftsanspruch von Wissenschaftspolitik und Öffentlichkeit und der Forcierung des Wettbewerbs im Wissenschaftssystem. |¹ Mit der Betonung des Wettbewerbsgedankens selbst wurde allerdings kein neues Prinzip eingeführt, vielmehr war der Wettbewerb um wissenschaftliche Erkenntnisse, um Lehrstühle, Schulbildung, Publikationen und Reputation schon immer ein wirksamer Mechanismus in der Forschung. Seit jeher gab es eine Differenzierung von Wissenschaftlerinnen und Wissenschaftlern nach Leistung, die jedoch nahezu ausschließlich in einer fachinternen Öffentlichkeit kommuniziert wurde. Neu ist, dass die Wissenschaftspolitik den fachinternen Wettbewerb in der öffentlich finanzierten Forschung für externe Steuerungsmaßnahmen

|¹ Vgl. Wissenschaftsrat: Empfehlungen zu Forschungsberichterstattung der Hochschulen, in: Wissenschaftsrat: Empfehlungen und Stellungnahmen 1980, Köln 1980, S. 39-45; Wissenschaftsrat: Empfehlungen zum Wettbewerb im deutschen Hochschulsystem, Köln 1985.

einsetzt und damit eine über die Fächer hinausgehende Offenlegung von Leistungsunterschieden zu schaffen versucht.

Die in den letzten Jahren eingeführten Bewertungs- und Steuerungsverfahren, die im Folgenden Gegenstand der Betrachtung sind, dienen zum einen der indikatorenbasierten Zuweisung von Ressourcen und zum anderen der Belohnung besonderer Leistungen. Grundsätzlich lassen sich dabei direkt steuernde von indirekt steuernden Verfahren unterscheiden. Während erstere durch bestimmte Vorgaben unmittelbar entsprechende Wirkungen erreichen wollen (wie z. B. die leistungsorientierte Mittelvergabe von Ländern an Hochschulen), zielen letztere über die Bereitstellung von Informationen und die Herstellung größerer Transparenz darauf, die Akteure zu eigenem Handeln im Sinne einer Qualitätsverbesserung oder eines bewussten Ressourceneinsatzes zu motivieren (z. B. Rankings oder Ratings). Die wichtigsten steuerungswirksam einsetzbaren Anreize sind dabei sowohl auf institutioneller als auch auf individueller Ebene Geld, Reputation, Gestaltungsfreiheit und Zeit. Die meisten in Deutschland verwendeten und hier betrachteten Verfahren zur Steuerung der Forschungsleistung setzen auf Bewertungsverfahren auf, da sie über Belohnungen wirken, für deren Vergabe es einer nachvollziehbaren Grundlage bedarf.

Implementiert wurden Bewertungs- und Steuerungsverfahren auf unterschiedlichen Handlungsebenen. Zu unterscheiden sind im Wesentlichen die vier Ebenen der Träger von Hochschulen und außeruniversitären Einrichtungen (Bund und Länder), der Forschung betreibenden Einrichtungen (Hochschulen und außeruniversitäre Forschungsinstitutionen), der intermediären Organisationseinheiten (Fakultät, Institute o. ä.) sowie der Wissenschaftlerinnen und Wissenschaftler innerhalb dieser Einrichtungen. Die Akteure auf diesen vier Ebenen sind freilich in unterschiedlichem Maße Anreize Setzende und von diesen Anreizen Betroffene.

A.II POSITIONSBESTIMMUNG

Die Gegensätzlichkeit der eingangs beschriebenen Positionen, die sich in Reaktion auf die Veränderungen in den Rahmenbedingungen der Forschung herauskristallisiert haben, erschwert die Entwicklung eines Konsenses. Der Wissenschaftsrat nimmt im Folgenden eine Position ein, welche die Unvereinbarkeit der Standpunkte anerkennt und einen Umgang damit möglich zu machen versucht, der beiden Seiten Zugeständnisse abverlangt.

Gute Forschung als Ziel der Steuerung

Bewertung und Steuerung von Forschungsleistung hat keinen Wert in sich selbst, sie soll gute Forschung ermöglichen und stärken. Für die Qualität von

11 Forschung gibt es zwar keine einfache Definition, aber ungeachtet disziplinärer Spezifika sind einige Grundsätze „guter“ Forschung eindeutig:

1 – Forschung ist eine Praxis eigener Art, eine Erkenntnispraxis, die zuerst der Logik der Wahrheitssuche folgt. Sie setzt Akteure voraus, die primär von Neugier und der Arbeit an ungelösten Problemen inspiriert sind und sich von der Freude an der Entdeckung des Neuen leiten lassen.

2 – Neben der Arbeit an gegebenen Themen und Problemen im Kontext vertrauter Theorien und Methoden zeichnet sich gute Forschung durch Originalität und Relevanz aus. Sie erfordert auch die Bereitschaft, sich selbst, die eigenen Annahmen, Theorien, Methoden und Daten in Frage zu stellen und ihre Bedeutung immer wieder neu zu prüfen.

3 – In ihren Ergebnissen und in der öffentlichen Kommunikation ist gute Forschung ihren eigenen theoretischen und methodischen Gütekriterien verpflichtet, aber offen für die Prüfung durch Dritte; sie dokumentiert und präsentiert Ergebnisse, die für andere Forscherinnen und Forscher nachvollziehbar und reproduzierbar sind, und sie ist in den Schlussfolgerungen transparent und verantwortlich gegenüber der Öffentlichkeit.

4 – Für gute Forschung sind Wahrhaftigkeit und gegenseitiges Vertrauen von grundlegender Bedeutung. Sie ist weder mit Täuschung noch mit geduldeten Nachlässigkeiten vereinbar. Die Gemeinschaft der Wissenschaftlerinnen und Wissenschaftler ist hierbei die erste Kontrollinstanz, zusätzlich ist die Rechtfertigung auch gegenüber der Öffentlichkeit eine selbstverständliche Pflicht. Die Suche nach Objektivität der Erkenntnis ist das zentrale regulative Prinzip. Sie impliziert neben disziplinspezifischen Kriterien der theoretischen und methodischen Arbeit eine wissenschaftliche Ethik, die sich zuerst den Regeln guter wissenschaftlicher Praxis verpflichtet.

Voraussetzungen guter Forschung

Gute Forschung erfordert nicht allein motivierte, kreative, risikobereite und theoretisch wie methodisch kompetente und reflektierte Forscherinnen und Forscher, sondern auch spezifische Voraussetzungen: eine arbeits- und kritikfähige, aber auch kritikbereite wissenschaftliche Gemeinschaft und eine Umwelt, die gute Forschung anerkennt und ermöglicht. Die langfristige Sicherung guter Forschung ist an bestimmte Voraussetzungen gebunden:

1 – Wesentlich ist die Anerkennung der besonderen Logik von Forschung. Diese muss nicht nur programmatisch, sondern institutionell so gesichert sein, dass nicht nur für die Wissenschaft, sondern auch für Politik und Gesellschaft, Kultur und Wirtschaft legitime Erwartungen und Grenzen, Rechte und Pflichten, Praktiken und Strukturen transparent und nachvollziehbar sind; gute Forschung braucht Vertrauen durch ihre Umwelt, Freiräume in ihrer alltägli-

chen Arbeit und die Chance, ihr eigenes Arbeits- und Zeitmuster zu praktizieren.

2 – Anerkennung und Vertrauen sollten sich in der Gestaltung von Rahmenbedingungen niederschlagen: in der Autonomie der Institution, auch als Erwartbarkeit des Handelns aller Beteiligten, z. B. für die Formen und die Dauer der Beschäftigung von Wissenschaftlerinnen und Wissenschaftlern; in rechtlichen Rahmenbedingungen, durch die Ergebnisse der Forschung in der Publikation und Verwertung geschützt werden; in einer Mittelausstattung, die den kontinuierlichen Grundbedarf in den Hochschulen ebenso sichert wie sie die Risikobereitschaft unterstützt, und die die Vielfalt der Erkenntnispraxis fördert.

3 – Ein breites Fundament qualitativ hochwertiger wissenschaftlicher Arbeit ist unverzichtbar und Basis für Spitzenleistungen („Exzellenz“) im Wissenschaftssystem. Die Unterscheidung von Grundlagenforschung, angewandter Forschung und forschungsorientiertem Service beschreibt nicht Wissenswelten unterschiedlicher Wertigkeit, sondern die nicht trennscharf unterscheidbaren Dimensionen der wissenschaftlichen Praxis insgesamt.

4 – Wissenschaftlerinnen und Wissenschaftler sollten bereit sein, nicht nur mit den Ressourcen, die ihnen Staat und Gesellschaft bereitstellen, verantwortungsvoll und effektiv umzugehen, sondern auch die Risiken der Arbeit an offenen und ungelösten Problemen zu suchen und diese gegenüber der Öffentlichkeit zu kommunizieren. Dies setzt u. a. voraus, nicht nur in den theoretisch oder methodisch vertrauten Bahnen zu arbeiten und Erfolge zu versprechen, sondern auch die Widerlegung von Prämissen und Erkenntnissen zu suchen sowie offensiv zu vertreten, dass Misserfolg und Scheitern genauso wie Fehlerhaftigkeit der eigenen Anstrengungen Teil der Forschungspraxis sind.

Externe Erwartungen an gute Forschung und wissenschaftliche Praxis

1 – Autonomie der Forschung bedeutet nicht Autarkie, etwa eine Verortung in Welten jenseits der gesellschaftlichen, kulturellen, wirtschaftlichen und politischen Wirklichkeit oder Finanzierbarkeit. Vielmehr setzt Wissenschaft immer einen offenen Austausch mit den öffentlichen, politischen, professionellen und gesellschaftlichen Erwartungen voraus, und dies nicht nur als Gegenleistung für die Tatsache, dass Wissenschaft auf Ressourcen von außen angewiesen ist. Wenn sich die Wissenschaftsgemeinschaft der Tatsache bewusst bleibt, dass sie in gesellschaftliche Strukturen und Erwartungen eingebunden ist, bedeutet dies weder die Unterwerfung unter den Markt für Waren und Güter noch die Ökonomisierung wissenschaftlicher Arbeit. Erwartungen an die Effizienz und Effektivität des Mitteleinsatzes sind deshalb genauso legitim wie die Erwartung der Relevanz der Themen sowie die Hoffnung auf Erfolg in der wissenschaftlichen Arbeit und in der Erschließung neuer, zukunftssträchtiger Forschungsfelder.

2 – Wissenschaft findet außerhalb des viel beschworenen Elfenbeinturms notwendige und unausweichliche Herausforderungen ihrer Kreativität, Beratungsfähigkeit und Problemlösungskapazität. Eine nicht allein innerwissenschaftliche sondern öffentliche Rechenschaftslegung ist deshalb der wissenschaftlichen Praxis nicht fremd, sondern inhärent. Außenerwartungen mit Blick auf wissenschaftlich bearbeitungsbedürftige Themen und Probleme sind bei Anerkennung der Eigenlogik der Forschung nicht illegitim, sondern notwendig und erwünscht.

3 – Gute Forschung profitiert von der umfassenden Bereitschaft, sich in der Lehre zu engagieren. Dies nicht nur aus Eigeninteresse, weil gute Forschung ohne guten wissenschaftlichen Nachwuchs nicht leben kann, oder wegen des gesellschaftlichen Qualifikationsbedarfs, sondern auch aus Gründen, die in der Forschung selbst liegen, zur Stärkung von Kritik und Selbstkritik. Hochschulische Lehre stellt immer auch die Herausforderung dar, sich dem unbefangenen Blick neuer Wissenschaftlergenerationen zu stellen und die Geltung und Relevanz der eigenen Befunde auch unter externen Erwartungen so zu zeigen und zu begründen, dass Gesellschaften über die Generationen hinweg lernfähig bleiben.

A.III ZIELE DER EMPFEHLUNGEN

Der Wissenschaftsrat erkennt die vorstehend skizzierte Koexistenz von legitimen Autonomieansprüchen der Wissenschaft und unabweisbaren externen Erwartungen an die Wissenschaft an. Daraus ergibt sich seine Aufforderung an Hochschulen und außeruniversitäre Einrichtungen, an die Wissenschaftspolitik in Bund und Ländern, an die Öffentlichkeit sowie an die einzelnen Wissenschaftlerinnen und Wissenschaftler, die Auseinandersetzungen um die Bewertung und Steuerung von Forschung zu versachlichen und eine Balance der unterschiedlichen Interessen anzustreben. Ziel der vorliegenden Empfehlungen ist es, Vorschläge zu machen, wie das gelingen kann.

Zu diesem Zweck wählt der Wissenschaftsrat folgendes Vorgehen: In einem ersten Schritt werden Beobachtungen zusammengetragen, die auf mögliche Zusammenhänge zwischen praktizierten Bewertungsverfahren und Steuerungsinstrumenten und positiven wie negativen Veränderungen in der Forschung hinweisen. In einem zweiten Schritt gilt es, diese Beobachtungen zu nutzen, um unter Anerkennung der bestehenden Interessenkonflikte Vorschläge zu entwickeln, wie positive Effekte verstärkt und negative Auswirkungen möglichst vermieden werden können. Erste Hinweise dazu wurden von anderer Seite

bereits gegeben und werden hier einbezogen. |² Ziel ist es, Bewertungsverfahren so zu gestalten, dass sie den Ansprüchen von Politik und Gesellschaft nach Rechenschaftslegung bei gleichzeitiger Wahrung der notwendigen Freiräume für die Forschung entsprechen und dabei Aufwand und Ertrag in ein angemessenes Verhältnis bringen. Die Spannung zwischen möglichst aussagekräftigen Forschungsbewertungen auf der einen Seite und möglichst geringem Aufwand sowie geringen unerwünschten Nebenwirkungen auf der anderen ist keineswegs die einzige nicht ohne weiteres auflösbare Zielspannung. Vielmehr ergeben sich solche auch zwischen der gegebenen Komplexität des Gegenstandes Forschung und der notwendigen Einfachheit der sie bewertenden und steuernden Verfahren sowie zwischen der Autonomie der Forschenden und dem legitimen Anspruch der Geldgeber auf Rechenschaft für die Verwendung öffentlicher Mittel. Ein weiteres Ziel der Empfehlungen des Wissenschaftsrates ist es deshalb, diese Zielspannungen bei der Anwendung von Steuerungsinstrumenten zuallererst bewusst zu machen und dort, wo notwendig und möglich, Korrekturen vorzuschlagen. Dabei ist der Wissenschaftsrat sich bewusst, dass er selbst Akteur auf dem Gebiet der Bewertung und Steuerung von Forschungsleistung und damit für die Erreichung dieser Ziele mit verantwortlich ist.

Der Fokus der Empfehlungen liegt auf der öffentlich finanzierten Forschung, unabhängig von ihrer jeweiligen Institutionalisierungs- und Organisationsform. Der Wissenschaftsrat wendet sich im Folgenden an Bund und Länder als Träger der Hochschulen und außeruniversitären Einrichtungen. Darüber hinaus adressiert er das Leitungspersonal der Hochschulen und der außeruniversitären Einrichtungen auf den unterschiedlichen Ebenen innerhalb dieser Einrichtungen, ferner richtet er sich aber auch an die Fachgemeinschaften als in unterschiedlicher Weise organisierte Vertretungen disziplinspezifischer Interessen. Nicht zuletzt werden die Wissenschaftlerinnen und Wissenschaftler selbst angesprochen.

|² Vgl. bspw. Jansen, D. (Hrsg.): Forschungspolitische Thesen der Forschergruppe "Governance der Forschung" - Rahmenbedingungen für eine leistungsfähige öffentlich finanzierte Forschung anlässlich der Tagung „Neue Governance für die Forschung“ Berlin, 14. und 15. März 2007, Deutsches Forschungsinstitut für öffentliche Verwaltung, Speyer 2007; Arbeitskreis der deutschen Universitätskanzler(innen) „Leistungsorientierte Mittelvergabe und Zielvereinbarungen“: Empfehlungen zur Gestaltung von Steuerungssystemen auf der Ebene Land/Hochschule, Gießen 2009; Robert Bosch-Stiftung: Thesenpapier der Gäste des 4. Berliner Wissenschaftsgesprächs der Robert Bosch-Stiftung, Stuttgart 2011.

B. Beobachtungen

Die Veränderungen im Wissenschaftssystem in den letzten beiden Jahrzehnten, zu denen auch die Einführung von Instrumenten der Bewertung und Steuerung von Forschungsleistung gehört, haben dazu beigetragen, Leistungsunterschiede in der Forschung transparenter und nachvollziehbarer zu machen. Während vorhandene Leistungsunterschiede zuvor meist nur einer eingeweihten Fachgemeinschaft bekannt waren, wurde im Prozess der Entwicklung von Bewertungsverfahren eine explizite Klärung des jeweiligen Qualitätsbegriffs in der Forschung erforderlich. Differenzen in der Leistungsfähigkeit sollten einer breiteren Öffentlichkeit offen gelegt werden. Eine an transparenten Kriterien orientierte Bewertung von Leistungen ist vergleichsweise weniger anfällig für Machtmissbrauch, sie ist strukturell offener und dynamischer, und sie fördert den Wettbewerb zwischen Institutionen und Individuen. Als weiterer Effekt zeichnet sich ab, dass viele Disziplinen – durch veränderte Publikationsstrategien, die Einwerbung wichtiger Drittmittelprojekte, die Bearbeitung auch gesellschaftlich relevanter Themen sowie durch Zusammenschlüsse von Forschenden – in den vergangenen zwei Dekaden deutlich sichtbarer geworden sind und sich im internationalen Wettbewerb behaupten können. Für Forschungseinrichtungen in Deutschland ist nicht nur eine größere Sichtbarkeit, sondern auch eine höhere Attraktivität durch Profilbildung und gezielte Förderprogramme zu beobachten. Zu befürchten steht jedoch, dass diese positiven Entwicklungen umschlagen können und zunehmend unerwünschte Nebenwirkungen eintreten.

Darüber, ob das Wissenschaftssystem in Deutschland inzwischen einen Zustand erreicht hat, an dem der Bogen gewissermaßen überspannt ist und zunächst möglicherweise positive Wirkungen durch negative Nebeneffekte in ihr Gegenteil verkehrt werden, besteht kein Konsens. Eine valide empirische Grundlage, die es ermöglichen würde, die Leistungsfähigkeit des Wissenschaftssystems vor Einführung dieser Instrumente mit derjenigen nach der Etablierung von Bewertungs- und Steuerungsverfahren zu vergleichen und dabei sonstige Veränderungen in demselben Betrachtungszeitraum – wie das Größenwachstum des Wissenschaftssystems, die relative Verknappung der Grundmittel, die Zunahme der Anforderungen an die Wissenschaft – auszuklammern, gibt es nicht. Bis-

lang liegen nur punktuelle Studien vor, die jedoch keine Kausalzusammenhänge nachweisen können. Auch der vergleichende Blick in andere Länder ist aufgrund der sehr unterschiedlichen Rahmenbedingungen nur bedingt hilfreich, und auch dort gibt es nur vereinzelte Studien zu den Wirkungen von Bewertungsverfahren und Steuerung in der Forschung. |³ Daher muss sich die nachfolgende Betrachtung auf Beobachtungen und plausible Annahmen zu möglichen Effekten und Reaktionen des Wissenschaftssystems auf die in Deutschland seit den 1990er Jahren eingeführten Verfahren der Bewertung und Steuerung von Forschungsleistung beschränken.

1.1 Zur Bewertung durch *peers*

Zur Bewertung wissenschaftlicher Leistung finden im Wissenschaftssystem unterschiedliche Begutachtungen durch *peers*, d. h. einschlägige Fachwissenschaftlerinnen und -wissenschaftler, in großer Zahl statt. Mit Blick auf die Forschung sind dies: Begutachtungen von Zeitschriftenartikeln, von Projektanträgen, von Wissenschaftlerinnen und Wissenschaftlern in Berufungsverfahren sowie von Einrichtungen in Form von Evaluationen. Ergänzt werden sie durch Verfahren mit Bezug auf die Lehre (v. a. Akkreditierungsverfahren). |⁴

Diese Leistungsbewertung durch *peers* in sogenannten *peer review*-Verfahren erlaubt im Vergleich der Bewertungsverfahren am ehesten ein fundiertes Qualitätsurteil über Forschung und eine institutionalisierte wissenschaftsinterne (Selbst-)Kontrolle. Sie wird der Vielfalt und Komplexität der Forschung am ehesten gerecht. *Peer reviews* werden auch deshalb immer öfter eingesetzt, weil sie wissenschaftspolitische Entscheidungen fundieren und legitimieren. Obgleich *peer review* wegen seiner klaren Vorteile als Bewertungsverfahren häufig eingefordert und durchgeführt wird, sind auch Nachteile evident: |⁵ In einem System permanenter und eng getakteter Bewertungsprozesse stößt das Begutachtungswesen an seine Grenzen, da der Aufwand für alle Beteiligten hoch ist:

_ Auf institutioneller Ebene wird der mit Begutachtungen verbundene Aufwand aufgrund mangelnder zeitlicher Abstimmung der verschiedenen Verfahren

|³ Einen Überblick über die Studien im internationalen Bereich liefert bspw. Butler, L.: Impacts of performance-based research funding systems: A review of concerns and the evidence, in: OECD: Performance-based funding for Public Research in Tertiary Education Institutions, Workshop Proceedings, Paris 2010, S. 127-165.

|⁴ Der Wissenschaftsrat wird voraussichtlich im Januar 2012 Empfehlungen zu Akkreditierungen als Element der Qualitätssicherung im Wissenschaftssystem abgeben.

|⁵ Zur Kritik am *peer review* vgl. bspw.: Osterloh, M.; Frey, B.: Das Peer Review-System auf dem ökonomischen Prüfstand, in: Kaube, J. (Hg.): Die Illusion der Exzellenz. Lebenslügen der Wissenschaftspolitik, Berlin 2009, S. 65-73.

sowie durch die häufig zu kurzen Intervalle zwischen den einzelnen Bewertungsverfahren als hoch empfunden. Problematisch ist überdies, dass die zur Unterstützung der *peer reviews* notwendigen internen wie externen Datenanfragen ebenso wenig standardisiert sind wie die von den Einrichtungen vorgehaltenen Daten.

- _ Auf individueller Ebene wird der Vorwurf der „Evaluitis“ in einem sich ständig beschleunigenden System erhoben. Dies wird begründet mit zeitlich und inhaltlich sich überlagernden, nicht abgestimmten Evaluationsverfahren, Datenerhebungen für Ratings und Rankings, Anfragen aus der Wissenschaftsforschung, Anforderungen verschiedener interner und externer Berichtssysteme sowie internen Qualitätssicherungsmaßnahmen. Für die einzelnen Forscherinnen und Forscher ist der Aufwand auch deshalb so groß, weil die Unterstützung durch eine serviceorientierte, wissenschaftsnahe Verwaltung nicht immer ausreichend gegeben ist. Für Begutachtungsaktivitäten scheint der Aufwand ungleich verteilt zu sein, d. h. es gibt eine kleine Gruppe besonders stark durch Projekt-, Institutionen-, Fächer- oder Artikelbegutachtungen belasteter Gutachterinnen und Gutachter. |⁶

Ein weiterer Kritikpunkt an der Begutachtung durch *peers* ist, dass sich die Urteile der *peers* zum Teil widersprechen oder nicht konsistent sind, dass gruppendynamische Prozesse zu einer Homogenisierung und damit einer innovationsfeindlichen Urteilsbildung führen können, dass Urteile zu sehr auf die Reputation von Personen und Institutionen abstellen, und dass die Gutachterinnen und Gutachter mitunter dazu neigen, ihre eigenen thematischen und methodischen Präferenzen in das Urteil einzubringen. Daher wird davor gewarnt, dass in *peer review*-Verfahren primär der wissenschaftliche Mainstream gefördert werde.

Ein häufig beschrittener Weg, mit diesen Problemen umzugehen, ist das „*informed peer review*“, bei dem das Urteil der *peers* durch ein bestimmtes Set an Indikatoren unterstützt wird. Dieses *informed peer review* ist sowohl einer reinen Reputationsbewertung als auch einer automatisierten indikatorenbasierten Bewertung vorzuziehen, da einerseits durch die Indikatorenbasis rein subjektive (Vor-) Urteile auf den Prüfstand gestellt werden und andererseits eine Ein-

|⁶ Zum Ausmaß der Begutachtungstätigkeiten auf Basis einer repräsentativen Umfrage zuletzt: Böhmer, S.; Neufeld, J.; Hinze, S. et al.: Wissenschaftler-Befragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten, iFQ-Working Paper No. 8, Bonn 2011, S. 131-135. Nur etwa ein Drittel der Befragten berichteten dort von einer mittleren oder hohen Begutachtungsaktivität. Die Gruppe der Hochaktiven wendet inzwischen allerdings durchschnittlich 13 % ihrer Arbeitszeit für Begutachtungstätigkeiten auf.

ordnung und mögliche Relativierung der für sich genommen nicht immer hinreichend aussagekräftigen Indikatoren durch Fachleute stattfindet.

Ein besonders wichtiges *peer review*-Verfahren ist das Berufungsverfahren, da es die zentrale Hürde für die Aufnahme in das Professorenamt ist. Seine Ausgestaltung wird im Rahmen der gesetzlichen Vorschriften weitestgehend den Fachbereichen bzw. Fakultäten überlassen, wobei die Präsidien der Hochschulen zunehmend stärkeren Einfluss nehmen. In der Begutachtung spielt neben der wissenschaftlichen Qualität die Passung zur Fakultät und zur anvisierten Ausrichtung der Professur eine große Rolle. Neben dem in den meisten Fällen üblichen Vortrag werden vor allem in der Vorauswahl der Bewerberinnen und Bewerber auch quantitative Indikatoren – zumeist die Zahl der Publikationen, gelegentlich aber auch bibliometrische Instrumente wie der „Impact-Faktor“⁷ von Zeitschriften, in denen die Bewerberinnen und Bewerber publizieren, oder der nach seinem Erfinder benannte Hirsch-Index⁸ – in das Qualitätsurteil einbezogen.

Verschiedentlich geäußerte Kritik am Begutachtungsprozess innerhalb der Berufungsverfahren bezieht sich auf die Engführung der Leistungsanforderungen, mangelnde Transparenz der Entscheidungsfindung gegenüber den Bewerberinnen und Bewerbern, die unzureichende Berücksichtigung fehlender Forschungszeit aufgrund familiären Engagements sowie auf die in Deutschland sehr lange Dauer der Verfahren.⁹ Dadurch ist – so die Kritik – nicht immer gewährleistet, dass die am besten qualifizierten Bewerberinnen bzw. Bewerber mit dem höchsten Potential in einem sachlichen und gegenüber den Betroffenen transparenten Verfahren berufen werden. Schwächen der Berufungsverfahren sind insbesondere deshalb nicht zu vernachlässigen, weil die Berufung eine Entscheidung von großer Tragweite für das gesamte Wissenschaftssystem ist, die innerhalb einer Einrichtung, aber auch für die jeweils betroffene Fachgemeinschaft für lange Zeit Weichen stellt.

⁷ Der Impact-Faktor einer Zeitschrift gibt an, wie häufig Artikel der Zeitschrift in einem begrenzten Zeitraum im Mittel zitiert werden. Er gilt als Maß für Ansehen und Bedeutung einer Zeitschrift.

⁸ Der Hirsch-Index, h-Index oder Hirsch-Faktor ist ein zitationsbasierter Indikator, der vor allem auf den Forschungsimpact als Qualitätsindikator zielt und dabei die Zitationshäufigkeit in der Breite des gesamten Publikationsoutputs einer Wissenschaftlerin bzw. eines Wissenschaftlers betrachtet; vgl. Hirsch, J. E.: An index to quantify an individual's scientific research output, S. 16569–16572, in: PNAS, November 15, 2005, vol. 102, no. 46. Die Interpretation und Bewertung dieses Indikators ist voraussetzungsvoller als die einfache Zahl annehmen lässt und bedarf u. a. einer Kenntnis des fachlichen Kontextes und des Lebensalters der jeweiligen Wissenschaftlerin bzw. des jeweiligen Wissenschaftlers.

⁹ Der Wissenschaftsrat hat im Jahr 2005 Empfehlungen gegeben, die auf eine Verkürzung und Qualitätssicherung des Verfahrens zielten, Wissenschaftsrat: Empfehlungen zur Ausgestaltung von Berufungsverfahren, Köln 2005; vgl. zu aktueller Kritik auch: Junge Akademie: Qualität statt Quantität – auch in Berufungsverfahren, Berlin 2010.

Quantitative Indikatoren werden in unterschiedlichen Verfahren der Bewertung und Steuerung eingesetzt, um durch Leistungsmessung entweder unmittelbar entscheidungsrelevante Informationen zu erzeugen oder aber eine Grundlage für informierte *peer review*-Prozesse zu liefern. Die populärsten Indikatoren zur Bemessung von Forschungsleistungen sind: Drittmittleinwerbungen, Promotionszahlen, Publikationszahlen und Zitationszahlen. Neben diesen einfacheren quantitativen Indikatoren erfahren auch komplexe Indikatoren mit Bezug zu Publikationen und Zitationen (bibliometrische Indikatoren) eine verstärkte Nachfrage. Beispiele dafür sind Zitationen pro Publikation im Vergleich zum Fachgebietsdurchschnitt oder der Hirsch-Index. Daneben werden weitere Indikatoren wie Patente, Lizenzeinnahmen, Preise, Vorträge oder Beiträge zur Organisation von Tagungen verwendet.

Die Verwendung quantitativer Indikatoren in der Bewertung von Forschungsleistungen dient in erster Linie der Komplexitätsreduktion und der Ermöglichung von Vergleichbarkeit. Quantitative Indikatoren erfüllen den Bedarf nach Messbarkeit und lassen sich vergleichsweise gut operationalisieren. Zudem kann durch eine formelbasierte Verrechnung quantitativer Indikatoren der Aufwand für Gutachterinnen und Gutachter reduziert werden. Allerdings kann die vermeintliche Einfachheit und Klarheit der Indikatoren zu einer leichtfertigen Verwendung in der Steuerung von Forschung führen. So berücksichtigen die gängigen Verfahren der Forschungsbewertung nicht hinreichend die selbstverstärkenden Prozesse beispielsweise im Hinblick auf Zitationszahlen und tragen damit zu einer Verschärfung des sogenannten „Matthäus-Effekts“ |¹⁰ bei. Die Verwendung von Indikatoren ist manipulationsanfällig und kann zu Zielverschiebungen führen, so dass schließlich nicht mehr die Verbesserung der Forschungsleistung das Ziel ist, sondern die quantitative Steigerung im Hinblick auf bestimmte Indikatoren.

Eine Überbetonung der Quantität gegenüber der Qualität in Bewertungsprozessen befördert eine Fokussierung auf solche Leistungen, die in irgendeiner Weise messbar sind: auf hohe Drittmittleinwerbungen, auf eine große Zahl an Publi-

|¹⁰ Der in der Wissenschaftssoziologie von Robert K. Merton beschriebene Matthäus-Effekt geht auf das biblische „Denn wer da hat, dem wird gegeben werden, dass er Fülle habe; wer aber nicht hat, von dem wird auch genommen, was er hat“ (Matthäusevangelium Kapitel 25, Vers 29) zurück. Merton beschrieb damit einen Effekt, nach dem wissenschaftliche Veröffentlichungen bekannter Autoren häufiger zitiert werden als Veröffentlichungen unbekannter Autoren. Inzwischen wird der Begriff auch für weitere Bereiche des Wissenschaftssystems verwendet, jüngst etwa für die Exzellenzinitiative, die in der Wahrnehmung einiger ihrer Kritiker in besonderem Maße befördere, dass ein kleiner Kreis von Universitäten durch diese verstärkenden Prozesse (ungerechtfertigterweise) sich noch weiter von den übrigen Hochschulen abkoppeln würde.

kationen, auf zahlreiche Promotionen etc. Bis zu einem gewissen Punkt kann eine Steigerung in diesen Messgrößen tatsächlich Forschungsleistungen indizieren und es ist nicht nachgewiesen, dass ein Fokus auf Quantitäten zwangsläufig einen Verlust an Qualität nach sich zieht. Dass jedoch eine Leistungssteigerung in messbaren Größen an (fachspezifische) Grenzen stößt und Sättigungseffekte eintreten können, sind berechnete Befürchtungen. Die Höhe der Drittmittel etwa kann Grenzen überschreiten, oberhalb derer der Aufwand für die Einwerbung und Verwendung der Drittmittel den möglichen Ertrag für die Forschung übersteigt – dies legen Studien für einige Teildisziplinen bereits nahe. |¹¹ Ebenso plausibel ist mit Blick auf Betreuungsrelationen, dass ab einer bestimmten (fachspezifischen) Zahl an Doktorandinnen und Doktoranden pro Professorin bzw. Professor eine angemessene Betreuung und Begleitung des wissenschaftlichen Nachwuchses nicht mehr möglich ist.

Auch eine allein auf Publikationszahlen ausgerichtete Publikationskultur kann Probleme erzeugen. Zwar ist eine Zunahme von Publikationsaktivitäten für die Forschung *prima facie* nicht negativ, da sie Ausweis des Größenwachstums und der Ausdifferenzierung der Wissenschaft sowie Ausdruck einer sinnvollen Strategie der wirkungsvollen Erkenntnisvermittlung sein kann. |¹² Soweit allerdings die Zahl an Publikationen und Zitationen den Ausschlag geben für die Positionierung bei der hochschulinternen leistungsorientierten Mittelvergabe, für Berufungsverfahren, Projektanträgen, Rankings sowie für Evaluationsverfahren, wird eine strategische Publikationspraxis implizit befördert, welche auf die Verwertung einer Erkenntnis in möglichst vielen Einzelpublikationen abzielt. In einigen Disziplinen ist eine Zunahme von unzulänglich qualitätsgeprüften Sammelbandveröffentlichungen, vor allem zur Dokumentation der Ergebnisse von Drittmittelprojekten und Tagungen zu beobachten, die dadurch unterstützt wird, dass die Drittmittelgeber auf solche Publikationen explizit Wert legen. Es ist anzunehmen, dass derartige Veränderungen der Publikationspraxis zusätzlich dadurch befördert werden, dass wissenschaftliche Verlage ein ökonomisches Interesse daran haben, möglichst viel zu publizieren.

Da in Bewertungsverfahren neben der reinen Zahl an Publikationen häufig auch die Zitationshäufigkeit von Publikationen eine Rolle spielt, ist in einigen Disziplinen eine Tendenz zur Ausrichtung ihrer Publikationsaktivitäten auf

|¹¹ Vgl. Jansen, D.; Wald, A.; Franke, K. et al.: Drittmittel als Performanzindikator der wissenschaftlichen Forschung, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, Jg. 59, Heft 1, 2007, S. 125-149.

|¹² Im Sinne des wissenschaftlichen Fortschritts kann es durchaus sinnvoll sein, über ein Forschungsprojekt möglichst früh und sukzessive zu publizieren; ebenso kann es zweckmäßig sein, über eine Steigerung des Publikationsoutputs die Aufmerksamkeit für ein Forschungsfeld zu erregen oder die Sichtbarkeit einer ganzen Disziplin zu erhöhen. Auch kann es sinnvoll sein, in einer Veröffentlichung nicht mehr als eine zentrale Erkenntnis unterzubringen, die so besser vermittelt und wahrgenommen werden kann.

Zeitschriften mit einem hohen „Impact-Faktor“ zu beobachten. Durch derartige Anpassungen, die fachspezifisch unterschiedlich ausgeprägt sind, kann es zu einer Zielverschiebung kommen, nach der Publikationsstrategien nicht mehr primär auf die fachspezifisch sinnvollste Vermittlung von Erkenntnissen und Forschungsergebnissen, sondern auf die Erfüllung bestimmter Kriterien der Leistungsbewertung hin optimiert werden. Ein Gegengewicht zu stark auf Quantitäten ausgerichteten Publikationskulturen setzt die DFG, indem sie in ihren Antragsverfahren seit 1. Juli 2010 an Stelle möglichst umfangreicher Publikationslisten nur noch eine begrenzte Zahl ausgewählter Publikationen zulässt. Welchen Effekt dies auf das auch von internationaler Praxis geprägte Publikationsverhalten tatsächlich haben wird, bleibt abzuwarten.

1.3 Zu Evaluationsverfahren

Evaluationen in der Forschung sind in der Regel auf Forschungsprojekte, auf Forschungseinheiten oder auf ganze Institutionen gerichtet und können die vorhandene Leistung einer Institution und ihre Zielerreichung im Nachhinein (*ex post*) oder mit Blick auf die Zukunft bewerten (*ex ante*).¹³ Sie dienen sowohl der Offenlegung und Bewertung von Leistungen als auch der Bestimmung von Entwicklungsperspektiven und Verbesserungspotentialen. Üblicherweise werden in Deutschland Evaluationen im Verfahren des *peer review* durchgeführt, d. h. es erfolgt eine Begutachtung durch für die Einrichtung oder das Forschungsprojekt einschlägige Fachwissenschaftlerinnen und -wissenschaftler. Die oben beschriebenen Nachteile von *peer reviews*, vor allem der hohe Aufwand und Abstimmungsbedarf, gelten somit auch für Evaluationsverfahren.

An Hochschulen sind Evaluationen von Forschung und/oder Lehre als Bestandteil des internen Qualitätsmanagements ein Instrument der Qualitätssicherung. Die genaue Ausgestaltung der Evaluation (Intervalle, Modus, etc.) ist in den einzelnen Hochschulgesetzen nicht vorgegeben und daher den Hochschulen überlassen. Evaluationen wirken unter Umständen eher indirekt steuernd und stoßen einen Prozess der Selbstvergewisserung an, aus dem Lerneffekte für die Weiterentwicklung und Verbesserung von Einrichtungen resultieren können. Kritiker beanstanden, dass nicht alle Verfahren tatsächlich Konsequenzen haben und Evaluationen vor allem an Hochschulen teilweise noch ein reines Ritual seien, das allein der formalen Erfüllung von Qualitätssicherungsvorgaben diene. Weitreichende direkte Konsequenzen wie etwa die Schließung von Hochschulinstituten oder einzelnen Lehrstühlen aufgrund mangelnder Leistun-

¹³ Beispiele für *ex-post*-Evaluationen sind die durch den Wissenschaftsrat durchgeführten Evaluationen außeruniversitärer Forschungseinrichtungen, Beispiele für eine *ex-ante*-Evaluation sind die von DFG und Wissenschaftsrat durchgeführten Bewertungsverfahren im Rahmen der Exzellenzinitiative.

gen in der Forschung seien aufgrund der Bedeutung der jeweiligen Einrichtungen für die Lehre in der Regel nicht zu befürchten, nennenswerte Mittelzuwächse aufgrund finanzieller Restriktionen nicht möglich. In welchem Maße forschungsbezogene Evaluationsverfahren an Hochschulen tatsächlich unmittelbar entscheidungsrelevant sind oder zumindest als Instrument der indirekten Steuerung eingesetzt werden, ist nicht eindeutig zu bestimmen und variiert vermutlich je nach Hochschule und Land.

An außeruniversitären Einrichtungen werden Evaluationen breit eingesetzt; die Verfahren unterscheiden sich je nach Trägerorganisation. Die Mitgliedseinrichtungen der Leibniz-Gemeinschaft (WGL) werden mindestens alle sieben Jahre durch den „Senatsausschuss Evaluierung“ evaluiert. Ergänzend dazu führen die wissenschaftlichen Beiräte der Institute Audits zur kontinuierlichen Qualitätssicherung durch. Die Max-Planck-Gesellschaft (MPG) lässt ihre Institute alle zwei Jahre durch deren mit externen Gutachterinnen und Gutachtern besetzte Fachbeiräte bewerten. Im Rahmen einer erweiterten, strategischen Evaluation werden zudem alle sechs Jahre thematisch verwandte Institute in einem Forschungsfeld zusammengefasst und untereinander verglichen. Die MPG betont außerdem den Evaluationscharakter ihrer *ex-ante*-Begutachtung bei der Wissenschaftlerauswahl (Harnack-Prinzip), ähnlich den Berufungsverfahren an Hochschulen. In der Helmholtz-Gemeinschaft (HGF) werden die Forschungsprogramme alle fünf Jahre zentrenübergreifend von externen Expertinnen und Experten begutachtet. Daneben werden auch die Zentren selbst evaluiert. Die Fraunhofer-Gesellschaft (FhG) evaluiert ihre Institute etwa alle fünf Jahre in so genannten Technologieaudits, wobei externe Vertreterinnen und Vertreter aus Wirtschaft und Wissenschaft in die Begutachtung eingebunden werden. Kontinuierlich werden die Institute an den eingeworbenen Projektmitteln aus der Wirtschaft und der öffentlichen Forschungsförderung gemessen, die intern unmittelbare Auswirkung auf die Höhe der Zuweisung der Grundfinanzierung haben. Extern erfahren die Institute darüber hinaus eine Bewertung ihrer Tätigkeit am Markt.

Im außeruniversitären Bereich können Evaluationen oftmals unmittelbare und vielfach weit reichende Auswirkungen haben, sowohl in finanzieller Hinsicht als auch mit Blick auf die Weiterentwicklung oder gar Weiterführung einer Einrichtung (thematisch, personell, strukturell). Die Einführung der Evaluationsverfahren als Qualitätssicherungsmaßnahme wird im außeruniversitären Bereich überwiegend positiv gesehen. Kritisiert wird allerdings – ebenso wie im Hochschulbereich – der damit verbundene Aufwand.

I.4 Zu Ratings/Rankings

Ratings/Rankings sind Verfahren der vergleichenden Bewertung und Offenlegung von Leistungen in der Forschung. Als Instrumente der indirekten Steuerung sind sie, anders als etwa die frühere *Research Assessment Exercise* (RAE) im Vereinigten Königreich, in Deutschland nicht an die Vergabe von Mitteln gekoppelt. Seit den 1980er Jahren haben sich auch hierzulande unterschiedliche Rankings etabliert. Dazu gehören Forschungs- und Hochschulrankings unterschiedlicher Anbieter in einzelnen Fächern oder als Leistungsvergleich ganzer Einrichtungen (z. B. Hochschul- und Forschungsranking des Centrums für Hochschulentwicklung CHE, Handelsblatt-Ranking). In Reaktion auf bestehende Verfahren hat der Wissenschaftsrat im Jahr 2004 selbst Empfehlungen zu einem Forschungsrating ausgesprochen und seitdem in einer Pilotstudie ein entsprechendes Verfahren in mehreren Fächern erprobt. |¹⁴ Die Methodik der vergleichenden Bewertung variiert, in der Regel erfolgt eine Verrechnung von Indikatoren in Rangplätze (Ranking); im Forschungsrating des Wissenschaftsrates werden qualitative und quantitative Indikatoren durch *peers* eingeordnet („*informed peer review*“). |¹⁵ Fachübergreifende Rankings, in denen ganze Hochschulen anhand weniger Indikatoren – wie vor allem Publikationszahlen, Fördergelder, ggf. wichtige Preise – miteinander verglichen und in eine Rangfolge gebracht werden, werden in Deutschland nicht durchgeführt. Sie spielen aber im internationalen Bereich eine große Rolle (v. a. *Academic Ranking of World Universities*, auch kurz „Shanghai-Ranking“, und das *World University Ranking*) und beziehen auch die deutschen Hochschulen und teilweise auch die deutschen außeruniversitären Forschungseinrichtungen ein. Diese Rankings sind aufgrund mangelnder Differenziertheit der verwendeten Kriterien und problematischer Vergleichsgrundlagen kritisch zu sehen.

Als Verfahren der vergleichenden Erfassung und Bewertung von Forschungsleistungen haben Ratings/Rankings den positiven Effekt, Leistungsunterschiede offen zu legen und eine Vergleichsgrundlage über Institutionen und Individuen hinweg zu liefern, die den forschenden Einrichtungen und den Fachgemeinschaften Anhaltspunkte für ihre Positionierung im nationalen und ggf. auch im

| ¹⁴ Vgl. Wissenschaftsrat: Empfehlungen zur vergleichenden Forschungsbewertung in den Geisteswissenschaften (Drs. 10039-10), Köln 2010; Wissenschaftsrat: Pilotstudie Forschungsrating, Empfehlungen und Dokumentation, Köln 2008; Wissenschaftsrat: Empfehlungen zu Rankings im Wissenschaftssystem - Teil 1: Forschung, in: Wissenschaftsrat, Empfehlungen und Stellungnahmen 2004, Bd. I, Köln 2004, S. 159-220.

| ¹⁵ Vgl. Wissenschaftsrat: Empfehlungen zu Rankings im Wissenschaftssystem - Teil 1: Forschung, in: Wissenschaftsrat, Empfehlungen und Stellungnahmen 2004, Bd. I, Köln 2004, S. 159-220; IREG Observatory on Academic Ranking and Excellence: Berlin Principles on Ranking of Higher Education Institutions, Berlin 2006.

internationalen Bereich bietet. So dienen differenzierte, wissenschaftsgeleitete Verfahren als Orientierungshilfe. Sie können eine wettbewerbliche Wirkung entfalten, welche auch die Strategien ganzer Länder, einzelner Hochschulen und Individuen beeinflussen kann. Problematisch ist jedoch, dass hinsichtlich der Aussagefähigkeit von Ratings/Rankings nicht immer hinreichend differenziert wird und so vorschnelle Urteile über die nationale und internationale Wettbewerbsfähigkeit erfolgen. Wie bei der direkten Bewertung mittels quantitativer Indikatoren kann es deshalb zu Zielverschiebungen kommen, so dass im Extremfall das Interesse an Inhalten und Qualität der Forschung durch das Interesse an einer Positionierung im Ranking verdrängt wird.

1.5 Zur Steuerung über Mittelallokation

Leistungsorientierte Zuweisung der Landesmittel an die Hochschulen |¹⁶

Verfahren der leistungsorientierten Mittelvergabe (LOM) der Hochschulbudgets der Länder an ihre Hochschulen dienen dazu, den Wechsel von der ministeriellen Detailsteuerung zu mehr Autonomie der einzelnen Einrichtungen bei gleichzeitiger Betonung von Qualität und Wettbewerb zu begleiten. Eine Mehrheit der Länder hat seit Mitte der 1990er Jahre unterschiedliche LOM-Modelle eingeführt und in den Folgejahren weiterentwickelt. Mit der Einführung von LOM-Systemen auf Länderebene waren verschiedene Intentionen verbunden: (1) Beseitigung der Zufälligkeiten eines historisch gewachsenen Finanzierungssystems, (2) Mittelzuweisung anhand klarer Kriterien und wissenschaftspolitischer Zielsetzungen, (3) finanzielle Honorierung von Leistungs- und Belastungsunterschieden in Forschung und Lehre, (4) Rechenschaftslegung gegenüber den Länderparlamenten und (5) ein sachgerechter Umgang mit knappen Haushaltsmitteln.

Aktuell werden in den einzelnen Ländern sehr unterschiedliche Mittelanteile anhand von lehr- und forschungsbezogenen Indikatoren verteilt. Während bis vor kurzem in den meisten Ländern vor allem Sachmittel mittels LOM-Verfahren verteilt wurden, werden zunehmend – bedingt auch durch die Einführung von Globalbudgets – auch Personalmittel anhand von Indikatoren vergeben. Die Spanne reicht von weniger als fünf bis zu knapp 100 % der gesamten Hochschulmittel, wobei die Umverteilungseffekte in der Regel begrenzt werden. |¹⁷ Diese Deckelung der LOM-Mittel und die Tatsache, dass das Gesamt-

|¹⁶ Auf die leistungsorientierte Mittelzuweisung an die außeruniversitären Einrichtungen wird im Folgenden nicht Bezug genommen, da die einzelnen Organisationen über jeweils spezifische Mittelzuweisungsmechanismen verfügen, die eine ausführlichere Einzelbetrachtung erfordern würden.

|¹⁷ Eine vergleichende Studie zu den LOM-Systemen der Länder hat HIS zuletzt im Jahr 2004 vorgelegt: Leszczensky, M.; Orr, D.: Staatliche Hochschulfinanzierung durch indikatorengestützte Mittelverteilung.

budget eines Landes für seine Hochschulen in der Regel begrenzt ist und damit die Indikatoren nur die Verteilung zwischen den Hochschulen regeln, können zur Folge haben, dass einzelne Hochschulen trotz Leistungssteigerungen Kürzungen hinnehmen müssen, wenn andere Hochschulen noch höhere Steigerungen aufweisen. Die Erfolge der einen Hochschule werden somit durch die Erfolge einer anderen Hochschule relativiert.

Die der LOM zugrunde gelegten Indikatoren bilden unterschiedliche wissenschaftspolitische Zielsetzungen ab, von denen die Steigerung der Forschungsleistung nur eines ist. Wichtigste Faktoren sind in der Regel lehrbezogene Indikatoren wie Zahl der Studierenden in der Regelstudienzeit, Zahl der Studienanfänger oder Absolventenzahlen. Dies deshalb, weil die Länder gehalten sind, die Hochschulen entsprechend der Nachfrage der Studierenden zu finanzieren. Dadurch werden die Spielräume für die Steuerung über weitere Anreize erheblich eingeschränkt. Indikatoren der Forschungsleistung sind vor allem forschungsbezogene Drittmittel (eingeworbene oder verausgabte; teilweise gewichtet nach Herkunft) und Promotionszahlen. |¹⁸ Je nach wissenschaftspolitischer Zielsetzung werden außerdem häufig Indikatoren der Gleichstellung und der Internationalität einbezogen. Anfänglich komplexere Verteilungsformeln wurden in den meisten Ländern sukzessive vereinfacht.

Kritiker der LOM-Verfahren bemängeln, dass Leistung als Allokationskriterium nicht immer im Vordergrund steht, sondern dass es den Ländern oftmals primär darum geht, ihre Hochschulen auf einer nachvollziehbaren Grundlage entsprechend ihrer jeweiligen Belastungen in Lehre und Forschung auszustatten und arbeitsfähig zu erhalten. Diese Intentionen wurden nicht immer klar benannt, vielmehr wurde die Einführung der LOM-Systeme vorrangig mit Wettbewerb und dem Leistungsprinzip begründet. Kritisiert wird auch, dass die Entscheidung für die der Mittelallokation zugrunde gelegten Indikatoren sich primär nach dem Vorhandensein entsprechender quantitativer Daten richtet. Wichtige Erwägungen dazu, welche Vor- und Nachteile die Indikatoren jeweils haben, ob sie überhaupt geeignet sind, das Leistungsprofil angemessen abzubilden und Steuerungswirkung entfalten zu können in dem Sinne, dass die Hochschulen sie auch beeinflussen können, bleiben demgegenüber oft nachrangig.

Dokumentation und Analyse der Verfahren in 11 Bundesländern, HIS-Kurzinformation A 2/2004, Hannover 2004; eine Aktualisierung der Studie ist für Ende 2011 angekündigt.

| ¹⁸ Häufig sind es nur diese beiden Indikatoren, ggf. ergänzt um die Anzahl an Habilitationen, anhand derer die forschungsbezogene Mittelverteilung erfolgt. Teilweise werden weitere Indikatoren wie der Erfolg bei kompetitiven Forschungsförderprogrammen einbezogen. Seltener werden auch Publikationszahlen, Patente, Wirtschaftskooperationen und Transferleistungen zugrunde gelegt.

Ein wesentlicher positiver und intendierter Effekt der LOM ist, dass Spielräume für eine diskretionäre Einflussnahme auf die Mittelzuweisung seitens der Länder eingeschränkt wurden, was einen Zugewinn an Rationalität und Egalität zur Folge hat. Es erscheint plausibel, dass sich die Hochschulen an den vorgegebenen Indikatoren ausrichten und dass damit Anreize für eine Leistungsorientierung gesetzt werden, zumal der Blick ins Ausland nahe legt, dass es auch bei geringen Finanzierungsanteilen zu Anpassungsreaktionen kommen kann.¹⁹ Allerdings führen leistungsunabhängige strategische Entscheidungen der Politik nach wie vor zu einer Kompensation der durch die LOM eigentlich beförderten Finanzierungsunterschiede zwischen den Hochschulen. Daher bleiben insbesondere mit Blick auf eine Steigerung der Forschungsleistung Zweifel an der Steuerungswirkung der leistungsorientierten Mittelvergabe in ihrer jetzigen Form.

Leistungsorientierte Mittelallokation innerhalb der Hochschulen

Die meisten Hochschulen in Deutschland wenden auch intern Verfahren der leistungsorientierten Mittelvergabe an. Diese hochschulinterne LOM liegt in der Zuständigkeit und Verantwortung der Hochschul- und Fachbereichsleitungen. Die Höhe der Budgetanteile, die von der Hochschulleitung an die intermediäre Ebene (Fakultäten/Departements/Fachbereiche) und von diesen wiederum an die Institute/Lehrstühle/Professuren verteilt wird, variiert. Dabei können unterschiedliche Handlungsvarianten beobachtet werden, die von internen Zielvereinbarungen bis zur Feinsteuerung über hoch ausdifferenzierte Indikatorensysteme reichen. Erfahrungen legen nahe, dass sich Hochschulen ganz oder zumindest überwiegend an dem LOM-Modell des jeweiligen Landes orientieren, dessen Kennzahlen übernehmen und dabei nicht hinreichend fächerspezifisch differenzieren. ²⁰ Eine Bewertung der überwiegend quantitativen Kriterien

¹⁹ Die australischen Universitäten erhielten ca. 7,9 % ihrer Einnahmen anhand eines indikatorenbasierten Finanzierungsmodells zugewiesen, was nahezu die gesamte Grundfinanzierung ist. Dieser Anteil ist zwar nicht hoch, zusätzliche Anreize wirkten aber dadurch, dass diese Form der Mittelallokation für die Universitäten leichter beeinflussbar scheint als der nicht-grundfinanzierte Mittelanteil, dass die Reputationsgewinne hoch sind und dass Studierende ihre Studienortwahl entscheidend von den Forschungsleistungen der Hochschule abhängig machen. Da das Modell zu mehr als 50 % den Indikator Drittmittel belohnte, waren konkrete Reaktionen der Universitäten zu beobachten. Effekte auf das Gesamtsystem konnten für Australien nachgewiesen werden, vor allem eine Konzentration der Forschungsmittel, Veränderungen im Forschungsportfolio der individuellen Forscherinnen und Forscher sowie eine allmähliche Trennung von Forschung und Lehre. Zum australischen System gibt es verschiedene Studien, vgl. bspw. Gläser, J.; Lange, S.; Laudel, G.; Schimank, U.: Evaluationsbasierte Forschungsfinanzierung und ihre Folgen, in: Mayntz, R.; Neidhardt, F.; Weingart, P. et al. (Hgg.): Wissensproduktion und Wissenstransfer. Wissen im Spannungsfeld von Wissenschaft, Politik und Öffentlichkeit, Bielefeld 2008, S. 145-170.

²⁰ Vgl. dazu auch folgende Studie: Leszczensky, M.; Orr, D.: Staatliche Hochschulfinanzierung durch indikatorengestützte Mittelverteilung. Dokumentation und Analyse der Verfahren in 11 Bundesländern, HIS-

durch Fachkolleginnen und -kollegen (*peers*) erfolgt in der Regel nicht. Auch wenn auf Hochschulebene verhältnismäßig weniger Mittel per LOM verteilt werden als auf Landesebene, werden inzwischen vermehrt komplexere Indikatoren-systeme eingeführt, die eine differenzierte Zuweisung von Mitteln anhand individueller Forschungsleistungen vorzunehmen versuchen und sich an der jeweiligen Struktur- und Entwicklungsplanung der Hochschule orientieren. Dieser Trend zur Ausdifferenzierung der Modelle ist gegenläufig zu den LOM-Modellen auf Landesebene, die schrittweise vereinfacht wurden. Ein grundsätzliches Problem ist: Einerseits können in einem komplexeren Modell unterschiedliche Qualitätskriterien und Voraussetzungen verschiedener Fachbereiche berücksichtigt werden, andererseits wird für die Formulierung von Verteilungsschlüsseln eine Komplexitätsreduktion notwendig, etwa durch die Zuweisung von Punktzahlen zu den spezifischen Leistungskriterien, so dass in Frage steht, ob der betriebene Aufwand für die Verteilung der Mittel letztlich angemessen ist, gerade wenn nur verhältnismäßig geringe Mittel erfasst werden.

Orientieren sich die Hochschulen an den Modellen der Länder, die vor allem Zuwächse bei Absolventen, Promotionen und Drittmitteln belohnen, besteht hochschulintern die Gefahr der Bevorzugung von absolventenstarken sowie von drittmittelstarken Fakultäten, ohne dass tatsächlich fachliche Leistungs-differenzen offen gelegt würden. Die Wirkung der hochschulinternen LOM ist in einer hochschul- und fächerübergreifenden Perspektive schwer einzuschätzen. |²¹ Besonders starke Effekte werden LOM-Verfahren im Fachbereich Medizin zugeschrieben; das könnte auf die besondere Finanzierungssituation und auf die Organisationsstrukturen in der Medizin zurückzuführen sein sowie auf die Tatsache, dass hier bereits sehr früh LOM-Verfahren eingeführt wurden. |²²

Kurzinformation A 2/2004, Hannover 2004: Von den 43 Hochschulen, die in Ländern mit staatlichem Formelmodell angesiedelt waren, hatten sich 13 Hochschulen (30 %) bei der Entwicklung entsprechender hochschulinterner Verfahren in hohem Maße am landesseitig verwendeten Modell orientiert. Für weitere 24 Hochschulen (56 %) traf dies zumindest in Teilen zu.

|²¹ Der Arbeitskreis der Universitätskanzler(innen) „Leistungsorientierte Mittelverteilung und Zielvereinbarungen“ kam im Jahr 2006 zu dem Schluss, dass die Steuerungswirkung der hochschulinternen LOM aufgrund der meist geringen Budgetrelevanz sehr begrenzt sei. Vgl. dazu Arbeitskreis der deutschen Universitätskanzler(innen) „Leistungsorientierte Mittelvergabe und Zielvereinbarungen“: Hochschulinterne ziel- und leistungsorientierte Mittelvergabe. Eine Handreichung, Oldenburg 2006, S. 14; vgl. auch: Jaeger, M.: Wie wirksam sind leistungsorientierte Budgetierungsverfahren in deutschen Hochschulen?, in: Nickel, S.; Ziegele, F. (Hgg.): Bilanz und Perspektiven der leistungsorientierten Mittelverteilung. Analysen zur finanziellen Hochschulsteuerung, CHE-Arbeitspapier Nr. 111, Gütersloh 2008, S. 36-50, hier S. 45.

|²² Eine Studie zur „Governance der Hochschulmedizin – Intendierte und nicht intendierte Effekte dezentraler Anreizsysteme am Beispiel der fakultätsinternen leistungsorientierten Mittelvergabe (LOM) in der Medizin“ wird vom Institut für Forschungsinformation und Qualitätssicherung (iFQ) voraussichtlich im Jahr 2012 abgeschlossen und veröffentlicht.

Auch jenseits dieses disziplinären Sonderfalls ist anzunehmen, dass Wissenschaftlerinnen und Wissenschaftler auf die von der Leitungsebene gegebenen Anreize reagieren und sich beispielsweise um mehr Promovierende oder mehr Drittmittelinwerbungen bemühen. Fraglich bleibt allerdings, ob steigende Drittmittelinwerbungen primär auf die monetäre Anreizsetzung der LOM-Systeme zurückzuführen sind, oder vielmehr auf Finanzierungsnotwendigkeiten bestimmter Forschungsvorhaben, oder auch auf eine zunehmende Bedeutung von Drittmitteln für die Reputationszuweisung.

W-Besoldung

Auf individueller Ebene wurden mit der W-Besoldung seit dem Jahr 2004 neue Leistungsanreize eingeführt. Wissenschaftlerinnen und Wissenschaftler können zu ihrem gegenüber der alten C-Besoldung deutlich niedrigeren Grundgehalt nun unter anderem so genannte „besondere Leistungsbezüge“ für besondere Leistungen in Forschung und Lehre und weiteren Tätigkeitsbereichen erhalten. Entsprechende Regelungen zur Gewährung dieser Leistungsbezüge sind in Verordnungen der einzelnen Bundesländer und in Richtlinien zur Vergabe von Leistungsbezügen der jeweiligen Hochschulen getroffen worden. Kriterien sind mit Blick auf Forschungsleistungen beispielsweise Ergebnisse von Forschungsevaluationen oder quantitative Parameter wie Drittmittel, Publikationen oder Patente. Die Hochschulen haben einen gewissen Freiraum bei der Wahl der Kriterien und können daher Zulagen entsprechend ihrer jeweiligen strategischen Zielsetzungen gewähren, somit die W-Besoldung steuerungsrelevant einsetzen.

Grundsätzlich handelt es sich bei der W-Besoldung um ein komplexes Regelungsgeflecht auf mehreren Ebenen – Bund, Land und Hochschule –, das sich noch in Entwicklung befindet und auch rechtlich noch überprüft wird, was eine übergreifende Bewertung erschwert. Die Kritik an den Regelungen der W-Besoldung zur Belohnung besonderer Leistungen in der Forschung richtet sich sowohl auf die grundsätzliche Frage, ob durch individuelle monetäre Anreize eine nachhaltige Steuerung der Forschungsleistung zu erreichen ist, aber auch auf die konkrete, oftmals zu detaillierte Ausgestaltung der Kriterien zur Vergabe der entsprechenden Zulagen.

Drittmittelfinanzierung an Hochschulen

Im internationalen Vergleich verfügt Deutschland über eine sehr vielfältige Förderlandschaft. Die wichtigsten Förderer sind die Deutsche Forschungsgemeinschaft (DFG), das Bundesministerium für Bildung und Forschung (BMBF), die Europäische Union (EU), verschiedenste nationale Stiftungen und die Wirtschaft. Die Bandbreite der Förderinstrumente erlaubt die Durchführung individueller Forschungsprojekte etwa im DFG-Normalverfahren ebenso wie große

kooperative Verbundprojekte, wie sie beispielsweise vom BMBF und von der EU gefördert werden, und schließlich auch die Durchführung besonders risikobehafteter Projekte etwa im Reinhart-Koselleck-Programm der DFG. Aufgrund der Breite und Vielfalt der zur Verfügung stehenden Förderinstrumente können ganz verschiedene, aus den Diskussionen in der Wissenschaftsgemeinschaft entstandene Forschungsfragen mit unterschiedlichen methodischen und zeitlichen Anforderungen und thematischen Ausrichtungen mithilfe finanzieller Unterstützung durch Mittel Dritter bearbeitet werden.

Angesichts der seit Jahren stagnierenden Grundfinanzierung der Forschung an den Hochschulen spielt die Finanzierung über Drittmittel eine immer größere Rolle. |²³ Mit zunehmender – wenn auch fachspezifisch unterschiedlich ausgeprägter – Angewiesenheit auf Drittmittel verstärkt sich die Wirkung der mit ihr verbundenen Verteilungspraxis. Verschärft wird dies dadurch, dass Drittmittel als Qualitätsausweis in Bewertungsverfahren und auch in den meisten LOM-Systemen auf Landes- und Hochschulebene herangezogen werden, ihnen somit also für die Verteilung der Grundfinanzierung eine zusätzliche Bedeutung zukommt. Das denkbar ungünstigste Szenario, das sich aus dieser Situation entwickeln könnte, wäre, dass Drittmittelinwerbungen nicht mehr der Finanzierung spezifischer Forschungsfragen dienen, sondern ihren Zweck primär in ihrer Verwendung als zentraler Leistungsindikator in der leistungsbezogenen Mittelvergabe hätten. |²⁴

Bei aller Vielfalt der Drittmittellandschaft in Deutschland ist die Forschung mit Drittmitteln in der Breite durch bestimmte Spezifika charakterisiert, die sich angesichts der zunehmenden Bedeutung der Drittmittelfinanzierung unverhältnismäßig auf das gesamte Forschungssystem auswirken können. |²⁵ Kritiker weisen darauf hin, dass Drittmittel mit kurzen Laufzeiten Gelegenheitsstrukturen schaffen, die langfristig bewirken können, dass kleinteilige Forschungsfra-

|²³ Zum veränderten Verhältnis von Grund- und Drittmittelfinanzierung an deutschen Hochschulen vgl. jüngst den Bericht des Vorsitzenden des Wissenschaftsrates zu aktuellen Tendenzen im Wissenschaftssystem: Neuere Entwicklungen der Hochschulfinanzierung in Deutschland, Juli 2011: Während die Grundmittel der Hochschulen in Deutschland zwischen 1995 und 2008 um nur 6 % gestiegen sind, haben sich die Drittmittelausgaben der Hochschulen mehr als verdoppelt. Der Anteil der Drittmittel ist von 11 auf fast 20 % des Gesamtbudgets angewachsen. Sieht man nur auf die Forschungsfinanzierung, so hat sich das Verhältnis zwischen Grundmitteln und Drittmitteln dadurch deutlich verschoben.

|²⁴ In einer durch das iFQ jüngst durchgeführten Befragung gaben Wissenschaftlerinnen und Wissenschaftler selbst mehrheitlich an, dass die Möglichkeit zur Beantwortung von Forschungsfragen und der positive Reputationseffekt wichtigste Triebfedern seien, sich auf dem Drittmittelmarkt zu engagieren und nicht die Verwendung des Indikators Drittmittel in leistungsorientierten Mittelvergabeverfahren. Vgl. Böhmer, S.; Neufeld, J.; Hinze, S. et al.: Wissenschaftler-Befragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten, iFQ-Working Paper No. 8, Bonn 2011, S. 85-89.

|²⁵ Vgl. dazu bspw. Torka, M.: Die Projektförmigkeit der Forschung, Baden-Baden 2009.

gen präferiert werden. Angesichts einer zeitlich befristeten Finanzierung scheint es nahe liegend, vor allem solche Forschungsfragen zu bearbeiten, für die bereits Daten oder Studien vorliegen (sekundäranalytische Forschung). Der Zeitdruck der Projektforschung könnte aber auch – ggf. im Verbund mit der Entwicklung immer schnellerer Messtechniken – eine Entwicklung befördern, derzufolge für Forschungsfragen zwar Primärdaten erhoben und veröffentlicht werden, eine ausreichende Analyse der gewonnenen Daten mit dem Ziel der Theoriebildung aber nicht mehr stattfindet.

Als Folge der gestiegenen Bedeutung der Drittmittelfinanzierung führen Kritiker besonders häufig an, dass sie zu einer Fokussierung auf solche Themen führen könne, die dem wissenschaftlichen Mainstream zuzuordnen oder gerade *en vogue* sind. Dahinter steht die Annahme, dass die Gutachterinnen und Gutachter in den mit der Drittmittelförderung verbundenen Antrags- bzw. Begutachtungsverfahren diese Themen eher für bearbeitenswert halten und eine starke Orientierung an den Themensetzungen umfangreicher Forschungsförderprogramme stattfindet. Problematisch ist schließlich, wenn Wissenschaftlerinnen und Wissenschaftler unabhängig vom eigentlichen Forschungsinteresse solche Forschungsfragen bearbeiten, die eine hohe Drittmitteleinwerbungssumme bei geringem Antragsaufwand versprechen. Dies kann zu einer Präferenz für Projekte und Themen führen, die vor allem politisch gewollt und daher finanziell stark gefördert werden. Hinzu kommt: Die Landschaft der Drittmittelförderung in Deutschland erlaubt zwar die Bearbeitung einer großen Bandbreite an Einzelfragen durch einzelne Forscherinnen und Forscher, vor allem im Normalverfahren der DFG; Erfolge in der Einzelförderung erhalten jedoch zunehmend weniger institutionelle Anerkennung als Erfolge bei der Einwerbung größerer, sichtbarer und kooperativ durchgeführter Verbundprojekte, die überdies meist ein größeres Drittmittelvolumen aufweisen. Kritiker führen des Weiteren an, dass es zu einer Konzentration von Drittmitteln auf wenige Standorte kommen könne, und damit zu mittel- und langfristigen Veränderungen in der Institutionenlandschaft.

Bei aller Kritik an möglichen negativen Auswirkungen einer zu großen Angewiesenheit der Forschung auf Drittmittel gilt aber auch: Kompetitiv vergebene Drittmittel sind ein wesentliches und unverzichtbares Element der Qualitätssicherung durch wissenschaftlichen Wettbewerb und können – bis zu einer gewissen Grenze – die Forschungsleistung positiv beeinflussen.

1.6 Zu den Standards guter wissenschaftlicher Praxis

Die Bewertung und Steuerung von Forschungsleistung dient dazu, die Einhaltung der Standards guter wissenschaftlicher Praxis zu sichern. Aufgrund der großen Bedeutung, die Vertrauen als Grundlage wissenschaftlicher Kommunikationsprozesse hat, bedroht jede Vernachlässigung der Standards guter wis-

senschaftlicher Praxis das Gelingen von Wissenschaft. Angesichts des Wachstums des Wissenschaftssystems und der intensiven medialen und politischen Beobachtung, unter der wissenschaftliche Praxis inzwischen steht, ist es nicht überraschend, dass wissenschaftliches Fehlverhalten heute häufiger in der Öffentlichkeit problematisiert wird als in der Vergangenheit. Nicht immer äußert sich dieses Fehlverhalten jedoch in der Form eines skandalträchtigen Plagiats oder gravierender Datenfälschung; häufig ist unredliches Verhalten im Forschungsprozess nur schwer erkennbar und äußert sich etwa als methodisch unsaubere, hinter eingeführten Standards zurückbleibende Arbeit, Vernachlässigung schwer zu beschaffender Quellen, unrechtmäßige Autorschaft, Vernachlässigung von Zitationsstandards, Anführen nicht gelesener Referenzen, Gefälligkeits-Zitation potentieller *reviewer*. Vielfach herrscht der Eindruck vor, dass derartige Praktiken in dem durch Bewertungs- und Steuerungssysteme verschärften Zeit- und Konkurrenzdruck erkennbar an Verbreitung gewonnen haben. Es liegt nahe, dass schon die der Forschung inhärenten hohen Erwartungen an Qualität und Originalität verbunden mit hohem Zeitdruck solches Fehlverhalten befördern. Auch kann der zunehmende Wettbewerbsdruck in einer internationalisierten Wissenschaftslandschaft, bei der eine immer größere Zahl von Forschenden aus dem In- und Ausland im Wettbewerb um Reputation und Beschäftigungschancen steht, eine Ursache für solches Fehlverhalten sein.

1.7 Zu den Folgen für die Lehre

Forschung ist in ein breites Beziehungsgeflecht eingebettet, so dass Veränderungen in den Rahmenbedingungen der Forschungsförderung sich auch auf andere Bereiche des Wissenschaftssystems auswirken. So kann die Lehre von guter Forschung grundsätzlich profitieren, wenn durch die Einwerbung zusätzlicher Forschungsmittel eine Zunahme der für die Lehre zur Verfügung stehenden Ressourcen erreicht wird. Daneben entstehen aber auch negative Nebenfolgen für die Lehre vor allem dann, wenn Anreize für Forschung so gesetzt werden, dass sie unmittelbar zu Lasten der Lehre gehen. Dies ist etwa dann der Fall, wenn Forschungsfreiräume als Belohnung eingesetzt werden und ein vollständiger Rückzug aus der Lehre stattfindet, ohne dass umgekehrt Belohnungen (etwa in Form einer besseren Ausstattung oder in Form von „Lehrfreisemestern“) für ein stärkeres Engagement in der Lehre gewährt werden. |²⁶

|²⁶ Ein Lehrfreisemester soll es den Professorinnen und Professoren ermöglichen, sich intensiv mit methodisch-didaktischen Fragen des Lehrens und Lernens auseinander zu setzen, mit dem Ziel, ihre Lehrkompetenz zu verbessern oder neue Lehr- und Lernkonzepte zu entwickeln, vgl. Wissenschaftsrat: Empfehlungen zur Qualitätsverbesserung von Lehre und Studium, Köln 2008.

Zu bedenken ist schließlich auch, dass im Wertesystem der Wissenschaft Forschung mit ihrer nationalen und internationalen Strahlkraft einen höheren Stellenwert einnimmt als die lokal begrenzt wirkende Lehre. Solange durch Forschung mehr Reputation zu erwerben ist und diese mehr Wertschätzung generiert, nimmt Lehre einen untergeordneten Rang ein. Dieses im Wissenschaftssystem selbst verankerte Reputationsgefälle erschwert die Etablierung wirksamer ausgleichender Anreize für die Lehre.

1.8 Zu den wissenschaftspolitischen Rahmenbedingungen

In den letzten Dekaden wurden die Hochschulen in Deutschland mit zahlreichen neuen wissenschaftspolitischen Vorgaben hinsichtlich der Studien- und der *governance*-Strukturen, der rechtlichen Rahmenbedingungen und Zielvorgaben konfrontiert. Für die Umsetzung der Vorgaben war oft nur wenig Einführungszeit vorgesehen. Eine Reflexion der angestoßenen Prozesse war und ist kaum möglich, wenn binnen kurzer Zeitabstände neue Zielvorgaben gesetzt und wissenschaftspolitische Steuerungsmaßnahmen eingeführt bzw. verändert werden, die auch die strategische Handlungsfähigkeit der Hochschulen betreffen. Der Umsetzungsdruck wirkt sich auf die Fakultäts-, Fachbereichs- und schließlich Lehrstuhlebene aus. Häufig hat die möglichst rasche Umsetzung neuer politischer Ziele Vorrang vor einer behutsamen Einführung von Neuerungen. Eine längerfristige Anpassung an die veränderten Bedingungen, die auch Orientierungs- und Anreizwirkung entfalten könnte, ist entsprechend schwierig. Auch im außeruniversitären Sektor hat es tiefgreifende und nachhaltige Reformen |²⁷ gegeben, die Rahmenbedingungen sind hier jedoch vergleichsweise stabiler.

|²⁷ Beispielsweise die Einführung der Programmorientierten Förderung in der Helmholtz-Gemeinschaft und die Etablierung des Evaluationssystems der Leibniz-Gemeinschaft.

C. Empfehlungen

Die geschilderten Beobachtungen möglicher Auswirkungen der Bewertung und Steuerung von Forschungsleistung fördern verschiedene Zielspannungen zu Tage, die es für die Formulierung von Empfehlungen zu berücksichtigen gilt: Zielspannungen beispielsweise zwischen innerwissenschaftlicher Prioritätensetzung bei der Themenwahl und Berücksichtigung außerwissenschaftlicher Relevanzen, zwischen Breiten- und Spitzenförderung, vor allem aber zwischen Aufwandsreduktion der Bewertungsverfahren und Anerkennung der Komplexität des Gegenstandes sowie zwischen wissenschaftlichem Anspruch auf Freiheit der Forschung und öffentlichem Anspruch auf Rechenschaftslegung und Effizienz.

Aus derartigen Zielspannungen ergeben sich immer wieder Kontroversen, die eingangs bereits skizziert wurden und Anlass für die vorliegenden Empfehlungen sind. Ziel sollte der pragmatische Umgang mit den unauflösbaren Zielspannungen sein: Wenn auch keine dauerhaft gültigen Handlungsoptionen aufgezeigt werden können, so sollten doch Extrempositionen und starke Ausschläge in Richtung des einen oder des anderen Pols vermieden werden.

C.I ÜBERGEORDNETE EMPFEHLUNGEN

Als Leitlinien für die Entwicklung von Handlungsempfehlungen (C.II) gibt der Wissenschaftsrat folgende übergeordnete Empfehlungen:

1 – Selbstreflexivität: Organisationen im Wissenschaftssystem müssen über institutionelle Mechanismen (z. B. Berichtssysteme, Benchmarking-Verfahren) verfügen, um in hohem Maße über sich selbst informiert und auskunftsfähig zu sein, und zwar nach Außen im Sinne der Rechenschafts- und Offenlegung ihrer Leistungen sowie in der Binnenperspektive, um strategie-, entscheidungs-, lern- und weiterentwicklungsfähig zu sein. Dazu gehört insbesondere auch der Aufbau von Bewertungskompetenz innerhalb der außeruniversitären Forschungseinrichtungen und Hochschulen.

2 – Verfahrensüberprüfung: Bewertungsverfahren und Steuerungsinstrumente dürfen nicht zu einem Selbstzweck werden. Ihre Gestaltung setzt zwingend eine Reflexion ihrer Funktion, ihrer erwünschten wie unerwünschten Wirkungen, ihrer Reichweite und ihrer Auswirkungen auf andere Bereiche voraus. Eine regelmäßige Überprüfung der Verfahren und ihrer Wirksamkeit – wo möglich unter Einbeziehung der Wissenschaftsforschung – mindert das Risiko ihrer Verselbstständigung oder Zweckentfremdung.

3 – Aufwandsbegrenzung: Bei allen Ansprüchen an Komplexität, Vielfalt und Reflexivität der Verfahren muss das Bemühen um Aufwandsbegrenzung im Vordergrund stehen, um die Belastung aller Beteiligten zu minimieren. Eine Aufwandsbegrenzung kann durch verschiedene konkrete Maßnahmen (beispielsweise Koordinierung verschiedener Verfahren, Standardisierung der Datenerhebung) erfolgen und bedeutet nicht zwangsläufig eine unangemessene Reduktion der Komplexität.

4 – Angemessene Taktung: Eine Balance zwischen den Unplanbarkeiten des Forschungsprozesses einerseits und den Ansprüchen an regelmäßige Rechenschaftslegung und Planbarkeit andererseits kann durch eine Entschleunigung bzw. angemessene Rhythmisierung der Zeittakte von Bewertungsprozessen erreicht werden.

5 – Vermeidung einer „Tonnenideologie“: In der Bewertung und Steuerung von Forschungsleistung bedeutet ein Zuwachs an Quantität nicht zwangsläufig einen Zuwachs an Qualität. In den entsprechenden Verfahren muss daher anerkannt und berücksichtigt werden, dass in erster Linie die Qualität der Forschungsleistung zählt.

6 – Adäquanz: Bei der Entwicklung von Verfahren der Bewertung und Steuerung der Forschungsleistung müssen die Unterschiede in den „Missionen“ der Hochschulen und der verschiedenen außeruniversitären Forschungseinrichtungen berücksichtigt werden. Diese können eine unterschiedliche Ausrichtung der Forschung bedingen. Sie prägen das Aufgabenprofil einer Einrichtung, das neben Forschung auch Aufgaben in Lehre, Service, Beratung etc. umfassen kann.

7 – Einbeziehung aller Akteursebenen: Für die Entwicklung zweckmäßiger und akzeptierter Verfahren der Bewertung und Steuerung von Forschungsleistung ist die Einbeziehung aller beteiligten Akteure und eine verbesserte Kommunikation zwischen den Ebenen erforderlich. Insbesondere ist eine offene Mitwirkung der Fachgemeinschaften an der Weiterentwicklung der Verfahren unverzichtbar. Nur so können fachspezifische Besonderheiten und Ansprüche in einer Weise einbezogen werden, die Effizienz und Effektivität verbindet.

8 – Finanzielle Voraussetzungen: Die Grundfinanzierung der Hochschulen und außeruniversitären Forschungseinrichtungen muss es ihnen ermöglichen, ihre Kernaufgaben in Forschung und Lehre, Nachwuchsförderung sowie – aufgabenprofilabhängig – Serviceleistungen zu erfüllen. Nur eine angemessene Grundfinanzierung vermag Sicherheit und Langfristigkeit zu gewährleisten und damit auch „risikoreiche“, unkonventionelle und zeitaufwendige Forschung abzusichern sowie Forscherinnen und Forschern den Wechsel von Forschungslinien zu ermöglichen.

9 – Attraktivität sichern: Steuerungssysteme sollten die Gewinnung, Förderung und Eigenständigkeit wissenschaftlichen Nachwuchses ermöglichen und ihm an den Hochschulen wie an außeruniversitären Einrichtungen attraktive Perspektiven und Arbeitsbedingungen bieten.

10 – Berücksichtigung der Lehre: Um der Bevorzugung der Forschung gegenüber der Lehre im Wissenschaftssystem entgegenzuwirken, sind auf der individuellen Ebene sowie auf der Ebene der Hochschulen wirksame Anreize für die Lehre und Steuerungsmechanismen, die auf eine Balance der Anreize für Lehre und Forschung zielen, dringend erforderlich.

11 – Vielfältigere Steuerungszugänge: Steuerung in der Forschung sollte auf ein insgesamt möglichst breit gefächertes Repertoire an Instrumenten und nicht vorrangig auf monetäre Anreize und quantitative Leistungsmessung setzen, um Zielverschiebungen und andere ungewollte Fehlentwicklungen zu vermeiden.

C.II EINZELNE EMPFEHLUNGEN

II.1 Zum Einsatz komplementärer Steuerungszugänge

Um zu vermeiden, dass einzelne, insbesondere monetäre Steuerungszugänge unerwünschte Anpassungen hervorrufen, empfiehlt der Wissenschaftsrat die Nutzung vielfältiger Steuerungsinstrumente, vor allem die Gewährleistung angemessener Spielräume für autonome Selbststeuerung und zeitliche Flexibilität.

In der Wissenschaft sind Autonomie und zeitliche Flexibilität hohe Güter, deren Anreiz- und Steuerungswirkung vor allem auf Wissenschaftlerinnen und Wissenschaftler bislang vernachlässigt wird. Der Wissenschaftsrat empfiehlt, den Forschenden je nach Karrierestufe und individuellem Bedarf ein adäquates Maß an Autonomie zu gewähren: Wissenschaftlerinnen und Wissenschaftlern sollte die Möglichkeit zugestanden werden, für begrenzte Zeiträume flexibel Schwerpunkte auf Forschungs- oder Lehrtätigkeiten zu legen; um eine Entkopplung von Lehre und Forschung zu vermeiden, sollte auch bei einer Schwer-

punktsetzung auf die Lehre stets in gewissem Umfang Forschung betrieben werden und umgekehrt. |²⁸ Ausnahmen sollten lediglich kürzere Abschnitte in Form von Freisemestern für Forschung oder Lehre darstellen. Der Wissenschaftsrat empfiehlt daher, dass die Hochschulen neben dem etablierten Instrument der Forschungsfreisemester verstärkt auch von dem Instrument der Lehrfreisemester Gebrauch machen sollten; dies nicht nur im Sinne der Flexibilisierung, sondern auch, um die üblicherweise stärkere Anreizsetzung in der Forschung auszugleichen. |²⁹ Darüber hinaus sollten Hochschulen den Forschenden häufiger als bislang für begrenzte Zeiträume die Teilnahme an Forschungskollegs ermöglichen. Derart flexible Karrieremodelle könnten in der (internationalen) Konkurrenz der Hochschulen um Wissenschaftlerinnen und Wissenschaftler von Vorteil sein.

Da Berufungsverfahren den Eintritt in das Wissenschaftssystem mit langfristigen Folgen steuern, sollten Hochschulleitungen und Länder diese mit geeigneten Standards und transparent gestalten, so dass unterschiedliche Leistungsdimensionen (Forschung, Lehre, Transferleistungen) unter Berücksichtigung von Diversität und Gleichstellungsaspekten angemessen einbezogen werden können (siehe auch C.II.3). |³⁰

Auch Zielvereinbarungen können ein sinnvolles Instrument der ergebnisorientierten qualitativen Steuerung sein. Sie können auf unterschiedlichen Ebenen (Individuum, intermediäre Ebene, Einrichtung, Träger) eingesetzt werden. Voraussetzung ihres sinnvollen Einsatzes ist ein hohes Maß an Führungskompetenz für die Formulierung geeigneter, nicht zu kleinteiliger oder einseitig auf Quantitäten ausgerichteter Ziele.

Des Weiteren empfiehlt der Wissenschaftsrat den gezielten Einsatz des Monitorings von Forschungsleistungen und von Instrumenten der indirekten Steuerung auf den unterschiedlichen Ebenen des Wissenschaftssystems. Zu denken ist an die Offenlegung von Evaluationsergebnissen und leistungsbezogenen Daten, um die auf Reputation basierende Steuerung zu einem großen Teil dem

|²⁸ Vgl. auch Wissenschaftsrat: Empfehlungen zu einer lehrorientierten Reform der Personalstruktur an Universitäten, Köln 2007.

|²⁹ Lehrfreisemester wurden etwa an der TU München vergeben, gefördert aus Mitteln des Wettbewerbs exzellente Lehre des Stifterverbandes und der Kultusministerkonferenz, vgl. http://portal.mytum.de/studium-und-lehre/lehrpreise/freisemester_fuer_lehre.html v. 9.10.2011. Einen weiteren starken Anreiz für ein stärkeres Engagement in der Lehre setzen Bund und Länder mit ihrem Förderprogramm zur Qualität der Lehre, vgl. <http://www.bmbf.de/foerderungen/15440.php> v. 7.10.2011 oder <http://www.gwk-bonn.de/index.php?id=269> v. 10.11.2011.

|³⁰ Vgl. Wissenschaftsrat: Empfehlungen zur Ausgestaltung von Berufungsverfahren, Köln 2005; Kritik an der gängigen Praxis äußert beispielsweise auch die Junge Akademie: Qualität statt Quantität – auch in Berufungsverfahren, Berlin 2010.

Wissenschaftssystem selbst zu überlassen. Ratings und Rankings, die wissenschaftlichen Ansprüchen genügen (vgl. C.II.5), können dies auf institutioneller Ebene unterstützen. Um die Nachteile des Reputationswettbewerbs, vor allem die Bevorzugung erfahrener Wissenschaftlerinnen und Wissenschaftler, auszugleichen, sollten Nachwuchsforscherinnen und -forscher schon frühzeitig unabhängig und eigenständig Forschungsprojekte durchführen, eigenverantwortlich für deren Qualität zuständig sein und die Ergebnisse selbst publizieren können. Entsprechend den Regeln guter wissenschaftlicher Praxis ist dabei die eindeutige Kennzeichnung der Autorschaft von Nachwuchswissenschaftlerinnen und -wissenschaftlern unabdingbar.

Wissenschaftliches Fehlverhalten und unzureichende Leistungen sollten in erster Linie durch die *scientific community* selbst sanktioniert werden. Bei nachgewiesenem wissenschaftlichem Fehlverhalten, das im Kontext von Drittmittelprojekten erfolgt ist, sollten auch die Fördereinrichtungen – wie dies beispielsweise die DFG bereits seit längerem praktiziert – Sanktionsmöglichkeiten prüfen. Vorbeugend müssen bei der Gestaltung von Bewertungs- und Steuerungsverfahren mögliche Auswirkungen auf die Einhaltung von Standards guter wissenschaftlicher Praxis bedacht werden. Die Sanktionierung mangelnder Forschungsleistungen auf individueller Ebene sollte im Rahmen von Evaluationsverfahren auch an Hochschulen möglich sein.

Zusammenfassung der Empfehlungen zum Einsatz komplementärer Steuerungszugänge:

1 – Gewährung eines höheren Maßes individueller Autonomie für flexible Schwerpunktsetzung im Karriereverlauf von Wissenschaftlerinnen und Wissenschaftlern.

2 – Gestaltung von Berufungsverfahren als zentrale qualitative Eintrittsteuerung auf Basis eines breiten Leistungsbegriffs (Forschung, Lehre und Transferleistungen) und nachvollziehbarer Kriterien.

3 – Nutzung von Instrumenten der indirekten Steuerung im Reputationswettbewerb der Forschung durch Hochschulen, Bund und Länder prüfen.

II.2 Zur Verwendung von *peer review*

Peer review ist als wissenschaftsgeleitetes Bewertungsverfahren mit primär qualitativer Orientierung in der Bewertung und Steuerung von Forschungsleistung unverzichtbar. Zugleich sind *peer review*-Verfahren aber sehr aufwendig und binden in hohem Maße Ressourcen, weshalb ihre Verwendung immer gut begründet sein sollte. Wie für alle anderen Bewertungsverfahren gilt auch für *peer review*-Verfahren, dass sie fehleranfällig sind und Fehlsteuerungen nicht völlig ausschließen können. Eine Reflexion und Qualitätssicherung der Verfahren ist daher ebenso unerlässlich wie die Einhaltung zentraler Standards,

welche die Zuverlässigkeit und Validität der Ergebnisse der Bewertung gewährleisten. Das betrifft insbesondere die Auswahl der Gutachterinnen und Gutachter; besonders wichtig sind möglichst vielfältig und kontextabhängig auch international zusammengesetzte Gutachtergruppen, in die Nachwuchswissenschaftlerinnen und -wissenschaftler systematisch einbezogen werden sollten. Die Gutachtergruppen sollten für typische Probleme von *peer review*-Verfahren – wie die Bevorzugung des Mainstreams, gruppendynamische Prozesse, etc. – sensibilisiert sein.

Zusammenfassung der Empfehlungen zur Verwendung von *peer review*:

- 1 – Begründung der Durchführung eines aufwendigen *peer review*-Verfahrens.
- 2 – Einhaltung der Standards solcher Verfahren, vor allem sorgsame Auswahl der Gutachtergruppe; dabei Vielfalt verschiedener wissenschaftlicher Perspektiven und Karrierestufen gewährleisten.
- 3 – Hinreichende Information der Gutachtergruppe über Probleme und mögliche Risiken von *peer review*-Prozessen.

II.3 Zu Indikatoren für Forschungsleistung

Allgemeine Empfehlungen

Quantitative Indikatoren für Forschungsleistung haben binnen weniger Jahre eine enorme Bedeutung gewonnen. Einen besonderen Stellenwert haben drittmittelbezogene und bibliometrische Indikatoren. Für die Entwicklung und Verwendung von Indikatoren der Forschungsleistung empfiehlt der Wissenschaftsrat:

Die *scientific communities* sollten sich über fachinterne Qualitätsstandards verständigen und diese dokumentieren, damit sie in die Anwendung, Interpretation und Weiterentwicklung von Indikatoren eingehen können. Der Begriff der Forschungsleistungen sollte möglichst breit gefasst sein und neben Wissens- und Technologietransferleistungen auch Dienstleistungen für die Forschung einbeziehen, wie etwa Leistungen im Zusammenhang mit Forschungsinfrastrukturen, |³¹ Herausgeberschaften, Gremienaktivitäten und Gutachtertätigkeiten.

Der Wissenschaftsrat weist darauf hin, dass es bei der quantitativen Steigerung von Forschungsaktivitäten einen abnehmenden Grenznutzen gibt. Er gibt zu

|³¹ Wissenschaftsrat: Übergreifende Stellungnahme zu Informationsinfrastrukturen, in: Wissenschaftsrat: Empfehlungen zu Forschungsinfrastrukturen, Köln 2011, S. 5-70, hier S. 58.

bedenken, dass ein linearer Zusammenhang zwischen der Steigerung bei bestimmten Forschungsindikatoren und der Forschungsqualität ab einem gewissen – fachspezifisch unterschiedlichen – Punkt nicht mehr zu erwarten ist. Dieses gilt es bei der Bewertung von Forschungsleistungen zu berücksichtigen. So ist beispielsweise für Promotionszahlen anzunehmen, dass bei zunehmender Zahl der zu betreuenden Promotionen die Qualität der Betreuungsleistung ab einem gewissen Punkt wieder abnimmt. Auch bedeutet ein Mehr an Drittmitteln nicht zwangsläufig ein Mehr an Forschungsqualität.

Da Indikatoren Forschungsleistung lediglich anzeigen, ohne sie zu erklären, ist stets eine Interpretation durch qualifizierte Personen erforderlich. Nur dann sind Indikatoren eine sinnvolle, weithin akzeptierte Ergänzung zu qualitativen Informationen und subjektiven Urteilen. Daher sind ergänzende, vor allem methodische Informationen zu den verwendeten Indikatoren erforderlich. Neben den Gutachterinnen und Gutachtern sollten auch die Adressaten der Bewertung bzw. ihre Nutzer Interpretationshilfen erhalten.

Auf individueller Ebene sollten bei der Bewertung und Honorierung von Forschungsleistungen etwa im Zusammenhang mit Projektbegutachtungen, bei der Entscheidung über Zulagen in der W-Besoldung und insbesondere bei Berufungsverfahren qualitative Informationen und inhaltliche Gesichtspunkte (wichtigste Publikationen, Forschungs- und Lehrleistungen, Wissens- und Technologietransfer, Dienstleistungen für die Forschung) in den Mittelpunkt rücken. Entscheidungen in Berufungsverfahren sollten nicht allein oder überwiegend auf der Grundlage bibliometrischer Indikatoren (etwa Hirsch-Indices) getroffen werden. Eine Bewertung individueller Forschungsleistung ausschließlich auf Basis quantitativer Indikatoren sollte gänzlich ausgeschlossen sein. Bei jeder Art der individuellen Bewertung sollten die Voraussetzungen unterschiedlicher Lebensphasen und Karrierestufen angemessen berücksichtigt werden.

Bei der Optimierung des Einsatzes von Forschungsindikatoren gilt es schließlich zu bedenken, dass sie in einem unauflösbaren Zielkonflikt im Bemühen um Datensparsamkeit stehen. In diesem Sinne erscheint eine Beschränkung auf wenige robuste Indikatoren notwendig.

Empfehlungen zu zentralen Indikatoren |³²

– Bibliometrische Indikatoren: Diese sind zur Unterstützung informierter *peer review*-Verfahren durchaus sinnvoll. Gleichwohl sind Methodenkenntnis und eine kritische Reflexion von Chancen und Risiken der Bibliometrie unabdingbar. Der Wissenschaftsrat empfiehlt daher, bei der Verwendung bibliometri-

|³² Vgl. Wissenschaftsrat: Pilotstudie Forschungsrating. Empfehlungen und Dokumentation, Köln 2008.

scher Indikatoren den Gutachterinnen und Gutachtern Hinweise zur Methodik zu geben oder (zumindest punktuell) Personen mit entsprechenden Methodenkenntnissen einzubeziehen, um spezifische Probleme stärker beachten zu können. Die erforderliche Berücksichtigung fachspezifischer Publikationskulturen, -praxen und -strategien bei der Interpretation der bibliometrischen Indikatoren ist nur durch *peers* möglich; sie sollten daher immer nur in *peer review*-Prozessen angewandt werden.

- _ Publikationslisten: Publikationslisten geben Auskunft über Art und Inhalt der Veröffentlichungen sowie über Fachprofil und Publikationsstrategie. Sie sind ein guter Indikator für qualitativ fundierte Urteile von *peers*. Publikationslisten können auf wenige ausgewählte Publikationen beschränkt werden.
- _ Publikationszahlen: Dieser Indikator gibt vor allem Auskunft über die Quantität der Veröffentlichungen und damit einen Hinweis auf die Produktivität der Forschung. Publikationszahlen sind als Leistungsindikator problematisch, da sie einen Anreiz geben, mehr zu publizieren, als für die Vermittlung wissenschaftlicher Erkenntnis zwingend notwendig ist („Salamitaktik“). Je nach Publikationspraxis eines Faches sollte beispielsweise nach Veröffentlichungstypen differenziert werden; nach Möglichkeit sollte eine Einordnung in den Fachgebietsdurchschnitt erfolgen. Der Indikator sollte von weiteren – eher qualitativen – Indikatoren flankiert werden.
- _ Zitationszahlen: Diese geben Auskunft über den Rezeptionserfolg einer Publikation. Auch hier ist eine Einordnung in den Fachgebietsdurchschnitt erforderlich. Der Indikator kann jedoch manipuliert werden („Zitationskartelle“) und sagt weniger über die Qualität der zitierten Publikation als über die Sichtbarkeit und Reputation der Autoren aus.
- _ Höhe der Drittmittel: Der Indikator Drittmittel (Einwerbungen oder Verausgabungen) gibt Auskunft über die Antragsqualität und implizit über die bisherigen Forschungsleistungen der Antragsteller, nicht aber über die Qualität der mit den Drittmitteln aktuell erbrachten Forschung. Es sollte eine Gewichtung der Drittmittel relativ zum Fachgebietsdurchschnitt vorgenommen werden, um die unterschiedlichen Möglichkeiten der Einwerbung und die fachspezifische Drittmittelpraxis zu berücksichtigen. Überdies sollten Drittmittel anhand der Vergabepraxis gewichtet werden: Drittmittel, die in einem qualitativen Begutachtungsprozess kompetitiv eingeworben wurden, sollten höher gewichtet werden als solche, die in nicht-kompetitiven Verfahren vergeben wurden. Schließlich weist der Wissenschaftsrat darauf hin, dass bei der Bewertung der Drittmittelhöhe berücksichtigt werden muss, dass ein linearer Zusammenhang zwischen Quantität und Qualität ab einem gewissen, fachspezifisch unterschiedlichen und vom Aufgabenprofil abhängigen Punkt nicht mehr gegeben ist.

- _ Listen der Drittmittelprojekte: Listen von Drittmittelprojekten geben Auskunft über die Art der durchgeführten Projekte sowie über die Strategie der Drittmitteleinwerbung. Solche Listen sollten auf möglichst wenige ausgewählte Projekte beschränkt werden. Sie können von *peers* in ein qualitatives Urteil eingebunden werden und ergänzen sinnvoll quantitative Indikatoren zur Höhe der Drittmittel.
- _ Promotionszahlen/Habilitationszahlen: Diese geben Auskunft über den Umfang der Förderung des wissenschaftlichen Nachwuchses, nicht über die Qualität der Nachwuchsförderung. Als Indikatoren für Forschungsleistung sind sie aufgrund möglicher Zielverschiebungen problematisch. Vor allem in automatisierten Prozessen (bspw. LOM-Verfahren) sollte daher berücksichtigt werden, dass ab einem gewissen, disziplinspezifischen Punkt kein positiver Zusammenhang zwischen Promotionszahl und Qualität der Betreuungsleistung zu erwarten ist. Auch in *peer-review*-Prozessen sollten sie lediglich als Nachweis von Aktivitäten der Förderung des wissenschaftlichen Nachwuchses verwendet werden. |³³
- _ Forschungspreise, Auszeichnungen, Keynote- und Plenarvorträge: Solche Auflistungen geben Auskunft über die Anerkennung und Bewertung der Forschungsleistung durch Dritte sowie über die Reputation und Sichtbarkeit einer Wissenschaftlerin bzw. eines Wissenschaftlers. Nicht auszuschließen ist, dass bei der Verwendung dieses Indikators solche Forschung bevorzugt wird, die modische Themen behandelt.
- _ Patenzahl und Patentwertigkeit: Der Indikator Patenzahl zeigt an, dass innovative Forschung mit Blick auf Anwendbarkeit stattfindet, sollte aber mit Informationen über die Patentwertigkeit gewichtet werden und sagt für sich nichts über die tatsächliche Anwendung und somit den Erfolg des Patents aus. Er sollte deshalb am besten in Verbindung mit Informationen zur Verwertung der Patente verwendet werden.
- _ Verwertung von Patenten: Erträge aus Patenten sind *ein* Indikator für den Erfolg der auf Basis der Forschung entwickelten Produkte und geben insofern eher qualitative Auskunft über Forschungsleistung. Angesichts der vielfältigen Funktionen von Patenten sind die darauf basierenden Indikatoren nicht allein für die Bewertung der Anwendbarkeit von Forschung geeignet.

|³³ Vgl. dazu: Anforderungen an die Qualitätssicherung der Promotion – Positionspapier des Wissenschaftsrates (Drs. 1704-11), Halle 2011.

Zusammenfassung der Empfehlungen zu Indikatoren für Forschungsleistung:

1 – Wissenschaftsinterne Verständigung über Qualitätsstandards als Basis der Indikatorenentwicklung erforderlich; dabei Anwendung eines breiten Leistungsbegriffs unter Einbeziehung von Dienstleistungen für die Forschung.

2 – Berücksichtigung des abnehmenden Grenznutzens bei der Bewertung von quantitativen Indikatoren wie der Höhe der Drittmittelinwerbungen und Promotionszahlen.

3 – Interpretation von Indikatoren durch qualifizierte *peers*, um fachbezogene Einordnung unter Berücksichtigung der spezifischen Funktion und methodischen Konstruktion des Indikators zu gewährleisten. Unterstützung von Gutachtergruppen sowie von Adressaten der Begutachtung durch Interpretationshilfen.

4 – Bewertung von einzelnen Wissenschaftlerinnen und Wissenschaftlern auf Basis qualitätsorientierter Verfahren.

II.4 Zu Evaluationsverfahren

Bei regelmäßig stattfindenden institutionellen Evaluationen sollten diese – gleich ob im außeruniversitären oder universitären Bereich – in der Regel nur alle fünf bis zehn Jahre durchgeführt werden. Bei der Festlegung der Intervalle sollte die Berufungsdauer des Leitungspersonals berücksichtigt werden. Durch größere Evaluationsintervalle wird der Aufwand für die Begutachteten sowie die begutachtenden Personen und Institutionen begrenzt, zudem werden Momentaufnahmen vermieden, mittelfristige Tendenzen sichtbar, und die zwischen zwei Evaluationen erfolgten Veränderungen können sinnvoll bewertet werden. Die Auftraggeber von Evaluationen (Zuwendungsgeber in Bund und Ländern, Trägereinrichtungen) sollten – sofern nicht bereits geschehen – ihre Evaluierungspraxis entsprechend anpassen.

Evaluationen beauftragende und durchführende Instanzen haben eine Verantwortung, die vielfältigen, unterschiedlich getakteten Evaluationen im Sinne einer Entlastung der betroffenen Individuen und Institutionen besser abzustimmen als dies bislang der Fall ist. Aufgrund der Verfasstheit des deutschen Wissenschaftssystems (föderale Struktur, vielfältige Träger), stellt dies eine besondere Herausforderung dar.

Gerade weil Evaluationen unmittelbar steuerungswirksam eingesetzt werden können und ihre Ergebnisse in die Weiterentwicklung von Einrichtungen,

Programmen oder ähnlichem eingehen, muss großer Wert auf die Qualität des Verfahrens und die Zuverlässigkeit der vorgenommenen Bewertung gelegt werden. |³⁴ Die Auftraggeber von Evaluationen sollten Kriterien und Verfahrensweise frühzeitig transparent machen und regelmäßig auf ihre Entscheidungsrelevanz, ihre Wirksamkeit und nicht intendierte Effekte hin kritisch überprüfen und ggf. anpassen. Für die notwendige, möglichst zeitnahe Umsetzung der aus Evaluationsergebnissen abzuleitenden Handlungsoptionen ist auf der Ebene der Hochschul- bzw. Einrichtungsleitungen ein hohes Maß an Führungskompetenz erforderlich, das auch einen professionellen Umgang mit Misserfolgen möglich macht.

Evaluationen sollten immer im *informed peer review*-Verfahren auf Basis einer soliden Datengrundlage durchgeführt werden und die Lektüre ausgewählter Publikationen beinhalten. Grundsätzlich gelten die in Abschnitt C.II.2 zum *peer review* gegebenen Empfehlungen. Ebenso wie die Gutachtergruppe sollten die Kriterien der Evaluation jeweils dem zu evaluierenden Gegenstand mit seinem spezifischen Aufgabenprofil auch jenseits des Kernbereichs der Forschung angemessen sein und können daher nicht vereinheitlicht werden. Wesentlich ist eine ausgewogene Mischung qualitativer und quantitativer Indikatoren, die ebenso wie das Evaluierungsprozedere vorab festgelegt und veröffentlicht werden sollten. Die Gutachterinnen und Gutachter müssen zu Beginn des Verfahrens über die Kriterien der Evaluation und die Funktion der verwendeten Indikatoren informiert werden. Allen Betroffenen sollte so weit wie möglich die Chance zur Teilnahme am Verfahren eingeräumt werden. Die Ergebnisse der Evaluation und die daraus sich ergebenden Handlungsoptionen sollten veröffentlicht werden.

Zusammenfassung der Empfehlungen zu Evaluationsverfahren:

- 1 – Durchführung von institutionellen Evaluationen im Abstand von fünf bis zehn Jahren.
- 2 – Verbesserte zeitliche Abstimmung der Durchführung von Evaluationsverfahren.
- 3 – Keine Folgenlosigkeit von Evaluationen als Instrumente der direkten Steuerung; regelmäßige Verfahrensüberprüfung im Hinblick auf intendierte und nicht intendierte Wirksamkeit erforderlich.

|³⁴ Vgl. auch die „Grundsätze für erfolgreiche Evaluationen“ des Wissenschaftsrates: Wissenschaftsrat: Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates (Drs. 10296-10), Lübeck 2010, S. 13-16.

4 – Durchführung von Evaluationen stets als *informed peer review*. Anpassung des Bewertungsverfahrens (Gutachtergruppe, Bewertungskriterien usw.) an den jeweiligen Evaluierungsgegenstand. Gewährleistung größtmöglicher Verfahrenstransparenz und Partizipationsmöglichkeiten als Basis für Akzeptanz.

II.5 Zu Ratings/Rankings

Rankings können nur dann durch Transparenz und Vergleichbarkeit eine positive Wirkung im Wettbewerb entfalten, wenn sie bestimmte Anforderungen erfüllen, hinter denen die meisten der bekannten Rankings nach wie vor zurückbleiben. Daher bekräftigt der Wissenschaftsrat an dieser Stelle seine bereits im Zuge der Entwicklung und Erprobung seines Forschungsratings gegebenen zentralen Empfehlungen: |³⁵

- _ An der Konzeptionierung und Durchführung der Verfahren sollten die verschiedenen Fachgemeinschaften als *peers* beteiligt werden;
- _ die Bewertung sollte immer im Modus des *informed peer review* und stets mehrdimensional, auf der Basis eines umfassenden Leistungsbegriffs der Forschung erfolgen; methodische Standards sollten eingehalten und die Datengrundlage offen gelegt werden.

Der Wissenschaftsrat wird sich nach Abschluss der Pilotphase seines Forschungsratings – voraussichtlich im Jahr 2013 – zu Ratingverfahren äußern und weiterführende Empfehlungen geben.

II.6 Zu Verfahren der Mittelallokation

Kontinuitätsorientierte Hochschulfinanzierung

Die Hochschulen sollten so finanziert werden, dass sie die Durchführung guter Forschung, Lehre, Nachwuchsförderung sowie ggf. weiterer Kernaufgaben wie Transfer-, Service- bzw. Infrastrukturleistungen kontinuierlich, projekt- und programmunabhängig leisten können. Der Wissenschaftsrat empfiehlt zu diesem Zweck, die bereits bestehenden Finanzierungsmodi weiterzuentwickeln. Eine Rückkehr zum *status quo ante* der inputorientierten Mittelzuweisung ist aus Sicht des Wissenschaftsrates weder zielführend noch wünschenswert. Stattdessen sollte das Finanzierungssystem

|³⁵ Vgl. Wissenschaftsrat: Empfehlungen zu Rankings im Wissenschaftssystem – Teil 1: Forschung, in: Wissenschaftsrat: Empfehlungen und Stellungnahmen 2004, Bd. 1, Köln 2004, S. 159-220; Wissenschaftsrat: Pilotstudie Forschungsrating. Empfehlungen und Dokumentation, Köln 2008.

- _ nach den Grundsätzen der Transparenz, Gleichbehandlung und Nachvollziehbarkeit gestaltet sein,
- _ die Belastungen und Bedarfe der Hochschulen für die Erfüllung ihrer Kernaufgaben berücksichtigen und
- _ die Ausstattungsbedarfe der Hochschulen nicht allein nach Maßgabe der Nachfrage der Studierenden regeln, sondern zusätzlich dem Grundsatz der kontinuierlichen Sicherung guter Forschung innerhalb der Hochschulen Rechnung tragen.

Soweit indikatorenbasierte Mittelallokationsmechanismen für die Grundfinanzierung der Hochschulen eingesetzt werden, sollten die ihnen zugrunde liegenden wissenschaftspolitischen Zielsetzungen klar benannt und kommuniziert werden. Indikatoren, die nicht leistungsorientiert sind, sollten nicht unter dieser Bezeichnung geführt werden, um nicht den falschen Eindruck zu erwecken, dass schon die Absicherung der regulären Aufgabenerfüllung in Forschung und Lehre als besondere Leistung belohnt würde. Um Fehlsteuerungen zu vermeiden, sollte bei zentralen Indikatoren (wie Studierendenzahlen, Promotionszahlen und die Höhe der Drittmittel) der abnehmende Grenznutzen berücksichtigt werden (vgl. C.II.3).

Ferner empfiehlt der Wissenschaftsrat Bund und Ländern mittelfristig eine die tatsächlichen Kosten, also auch die indirekten Kosten berücksichtigende Finanzierung von Drittmittelprojekten. |³⁶ Da nach bisheriger Praxis Drittmittelprojekte eine teilweise Mitfinanzierung aus Grundmitteln benötigen (v. a. Infrastruktur, Verwaltung), würde erst eine Umstellung auf eine solche kostendeckende Finanzierung den Hochschulen die erforderlichen Freiräume für ihre strategische Weiterentwicklung verschaffen. Gleichwohl muss gesichert sein, dass diese auch indirekte Kosten abdeckende Projektfinanzierung nicht zu einer Minderung der Grundfinanzierung führt.

Auch innerhalb der Hochschulen sollte die Ermöglichung der kontinuierlichen Erfüllung der Kernaufgaben den Modus der Mittelzuweisung bestimmen. Auf dieser Ebene sollten zusätzlich fachspezifische Differenzierungen in der Mittelzuweisung vorgenommen werden. Vor allem wegen des Anspruchs an fachliche Differenzierung sind für die Weiterentwicklung der hochschulinternen Mittelverteilung neue Dialogforen zwischen den unterschiedlichen beteiligten bzw. betroffenen Ebenen – Hochschulleitung, Fakultäten bzw. Fachbereiche, Dekanate, einzelne Wissenschaftlerinnen und Wissenschaftler – unerlässlich.

| ³⁶ Vgl. dazu auch Wissenschaftsrat: Empfehlungen zur künftigen Rolle der Universitäten, Köln 2006, S. 55.

Jenseits der Grundfinanzierung sollten sowohl auf Ebene der Länder als auch auf Ebene der Hochschulen und Fakultäten Zusatzmittel für die Förderung bestimmter strategischer Ziele bereitgestellt werden. Wie bereits in einigen Ländern praktiziert, könnten diese Mittel mittels Zielvereinbarungen oder im Rahmen gezielter Förderprogramme antragsbasiert vergeben werden. Derartige Zusatzmittel können im Gegensatz zu den Grundmitteln durchaus diskretionär, nicht aber arbiträr vergeben werden.

Ziel der mit diesen Zusatzmitteln verbundenen Anreizsetzung muss nicht notwendigerweise die Förderung von Spitzenforschung sein: Während die Länder ihre Hochschulen beispielsweise für Erfolge in bestimmten strategisch gesetzten Bereichen prämiieren können, können die Hochschulen intern auch andere Steuerungsziele – etwa Profilbildung oder besonders unkonventionelle, risikoreiche, innovative Forschung – setzen, und diese mittels Prämien und Risikokapital fördern. Auch die Fakultäten haben durch diese zusätzlichen Mittel die Möglichkeit, spezifische Leistungen zu belohnen bzw. gezielt zu fördern.

Finanzierung außeruniversitärer Forschungseinrichtungen

Die außeruniversitären Forschungseinrichtungen sollten weiterhin auf Basis der mit den Zuwendungsgebern vereinbarten übergreifenden forschungspolitischen Ziele berechenbare finanzielle Perspektiven und verbesserte Rahmenbedingungen erhalten, damit sie auch künftig ihre jeweiligen missionsspezifischen Aufgaben erfüllen können.

Bewertungskompetenz für die W-Besoldung

Das gesamte System der W-Besoldung befindet sich derzeit in einer dynamischen Entwicklung und einer rechtlichen Überprüfung. Gleichwohl sieht der Wissenschaftsrat schon jetzt in den Hochschulen grundsätzlichen Bedarf, interne Bewertungskompetenz aufzubauen und weiterzuentwickeln, die für Entscheidungen über die Bepreisung von Leistungen und die Vergabe besonderer Leistungsbezüge unverzichtbar ist. Bei der Beurteilung individueller Leistungen ist von der ausschließlichen Verwendung quantitativer Indikatoren Abstand zu nehmen (vgl. C.II.3). Stattdessen sollten Instrumente qualitativer Selbstberichte stärker als bisher Verwendung finden. Zu bedenken ist, ob auch im Zusammenhang der W-Besoldung eine Bewertung in größeren zeitlichen Abständen angemessen wäre. Der Wissenschaftsrat behält sich vor, an geeigneter Stelle ausführlicher zur W-Besoldung Stellung zu nehmen.

Grundsätzlich weist der Wissenschaftsrat darauf hin, dass Zulagen für besondere Leistungen in der Forschung etwaige Fehlentscheidungen in Berufungsverfahren nicht nachträglich kompensieren können.

Leistungsanreize über Drittmittelfinanzierung

Drittmittel erfüllen eine zentrale wettbewerbliche Anreiz- und Qualitätssicherungsfunktion in der Forschung im hochschulischen wie im außeruniversitären Bereich. Sie sollten in ihrer Bedeutung für die Finanzierung der Forschung wieder auf eine ergänzende, nicht schon die Kernaufgaben der Hochschulen absichernde Funktion zurückgeführt werden. Die beschriebenen negativen Folgen einer zu stark auf Drittmittel angewiesenen Forschungsfinanzierung sollten eingeschränkt werden.

Die in Deutschland bestehende Vielfalt der Drittmittelförderung sollte unbedingt beibehalten werden; dazu gehören auch spezielle, auf Spitzenforschung ausgerichtete und im Antragsverfahren vergebene Forschungsmittel der Länder. Auch in Zukunft sollte eine große Bandbreite methodischer und thematischer Ansätze durch Drittmittel gefördert werden, noch stärker als bislang sollten dabei bedarfsabhängig variable Projektlaufzeiten möglich sein.

Zusammenfassung der Empfehlungen zu Verfahren der Mittelallokation:

- 1 – Grundfinanzierung der Hochschulen nach Maßgabe einer kontinuierlichen Ermöglichung guter Forschung, Lehre, Nachwuchsförderung und ggf. Infrastruktur- und Serviceleistungen.
- 2 – Ermittlung der Ausstattungsbedarfe nach Maßgabe der Kernaufgaben der Hochschulen – auch jenseits der Studierendennachfrage.
- 3 – Weiterentwicklung der indikatorenbasierten Finanzierung zu einem transparenten und nachvollziehbaren System; Vermeidung des „Etikettenschwindels“ der Leistungsorientierung; Einführung von Korridoren für die wichtigsten Indikatoren.
- 4 – Finanzierung von Bund-/Länder-geförderten Drittmittelprojekten entsprechend der tatsächlichen Kosten, d. h. inklusive der indirekten Kosten.
- 5 – Bereitstellung von Mitteln für die Förderung strategischer Ziele, sowohl auf Landesebene als auch hochschulintern.
- 6 – Aufbau interner Bewertungskompetenz an den Hochschulen für die W-Besoldung; Verzicht auf die Verwendung rein quantitativer Indikatoren für die Bewertung von Individuen.
- 7 – Erhalt und weiterer Ausbau der Vielfalt der Drittmittelförderung durch Bund und Länder als wettbewerblicher Leistungsanreiz.

Zur Verminderung des Aufwandes, der mit Verfahren der Bewertung und Steuerung von Forschungsleistung verbunden ist, können zwei Arten von Maßnahmen beitragen: erstens eine Reduzierung der Zahl von Evaluationen und ähnlichen Verfahren sowie zweitens eine Verbesserung und Effizienzsteigerung der kontinuierlichen Dokumentation von Forschungsleistungen, die der Wissenschaftsrat daher dringend empfiehlt. Eine solche Dokumentation ist zudem essentielle Grundlage einer institutionell abgesicherten Selbstreflexivität und Lernfähigkeit des Wissenschaftssystems.

Regelmäßige, umfassende und zugleich einfach zu handhabende Dokumentationen von Forschungsleistungen lohnen sich für alle beteiligten Akteure: Wissenschaftlerinnen und Wissenschaftler können Berichtsansforderungen seitens der Einrichtungsleitungen entsprechen, ohne dass wiederholte Datenanfragen stets zu hohem Bearbeitungsaufwand führen. Als Mitglieder von Fachgemeinschaften erhalten sie zugleich ein umfassendes Bild der eigenen Forschungsleistungen. Wissenschaftliche Einrichtungen können ihre Leistungen beobachten und auf informierter Grundlage, ggf. auch durch ein vergleichendes Benchmarking, strategische Entscheidungen hinsichtlich ihrer Weiterentwicklung und Profilbildung treffen. Schließlich erhält die Wissenschaftspolitik fundierte Informationen, um Entscheidungen über Ressourcenströme und wissenschaftspolitische Schwerpunktlegungen zu treffen sowie Erkenntnisse über die Effekte ihrer Maßnahmen zu gewinnen. Nicht zuletzt kann die Wissenschaftsforschung auf der Basis umfassender Datengrundlagen belastbarere Aussagen über beabsichtigte und unbeabsichtigte Effekte von Verfahren der Leistungsbewertung und Steuerung treffen. Insgesamt sieht der Wissenschaftsrat noch erheblichen Verbesserungsbedarf, um ein gutes Dokumentationswesen – auch als Grundlage eines funktionierenden Monitorings der eingeführten Verfahren – zu implementieren.

Der Wissenschaftsrat empfiehlt eine rasche, möglichst länderübergreifende Verständigung über Mindeststandards der Datenerhebung und eine Harmonisierung der vielerorts noch im Aufbau befindlichen Berichtssysteme an Hochschulen und außeruniversitären Einrichtungen. Ein Vorschlag, den die Steuerungsgruppe Forschungsrating des Wissenschaftsrates derzeit zur Standardisierung von Daten für die Forschungsbewertung erarbeitet, sollte als Ausgangspunkt weiterer Bemühungen dienen. |³⁷

Damit die Verbesserung der Berichtssysteme nicht zu Lasten der Forschenden geht, ist innerhalb der Einrichtungen sowie auch auf politischer Ebene eine

|³⁷ Mit der Veröffentlichung entsprechender Empfehlungen ist Anfang 2012 zu rechnen.

49 Unterstützung durch die Wissenschaftsadministration zwingend erforderlich. Diese sollte in enger Abstimmung mit der Wissenschaft die Berichtssysteme weiterentwickeln, pflegen, bei internen wie externen Anfragen Auskünfte geben können und die Umsetzung von Evaluationsergebnissen unterstützen. Die direkten und indirekten Kosten für Einrichtung und kontinuierliche Pflege der Dokumentationssysteme werden ausgeglichen durch den verringerten Aufwand in der Datenbereitstellung.

Zusammenfassung der Empfehlungen zum Dokumentationswesen:

1 – Verständigung über Mindeststandards der Datenerhebung und Harmonisierung der vorhandenen bzw. im Aufbau befindlichen Berichtssysteme der Hochschulen und außeruniversitären Einrichtungen.

2 – Entlastung der einzelnen Wissenschaftlerinnen und Wissenschaftler bei Datenvorhaltung, -bereitstellung und -pflege durch eine serviceorientierte, wissenschaftsnahe Administration.