

WISSENSCHAFTSPOLITIK
IM DIALOG

7/2013

Eine Schriftenreihe der
Berlin-Brandenburgischen
Akademie der Wissenschaften

DER DEUTSCHE SONDERWEG
IN DER MESSUNG VON FORSCHUNGSLEISTUNGEN

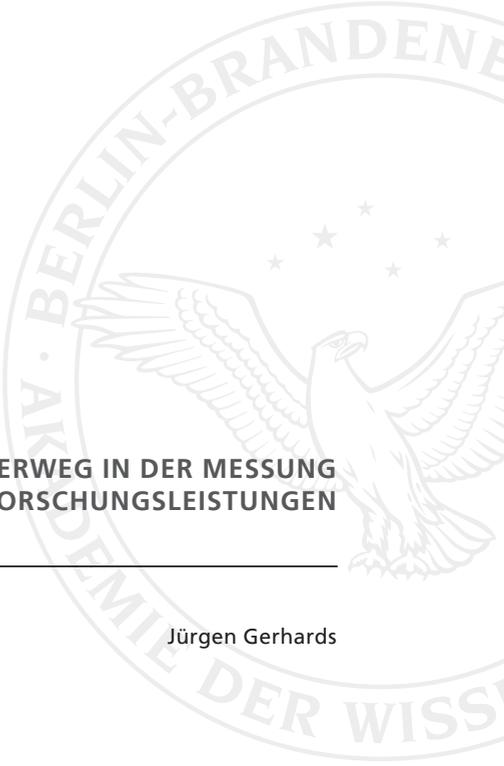
Jürgen Gerhards



berlin-brandenburgische
AKADEMIE DER WISSENSCHAFTEN

Berlin-Brandenburgische Akademie der Wissenschaften (BBAW)

DER DEUTSCHE SONDERWEG IN DER
MESSUNG VON FORSCHUNGSLEISTUNGEN



**DER DEUTSCHE SONDERWEG IN DER MESSUNG
VON FORSCHUNGSLEISTUNGEN**

Jürgen Gerhards

WISSENSCHAFTSPOLITIK
IM DIALOG

7/2013

Jürgen Gerhards (*1955) ist Professor für Soziologie an der Freien Universität Berlin und Mitglied der Berlin-Brandenburgischen Akademie der Wissenschaften.

Kontakt: j.gerhards@fu-berlin.de

Herausgeber: Der Präsident der Berlin-Brandenburgischen Akademie der Wissenschaften und die interdisziplinäre Arbeitsgruppe *Exzellenzinitiative* der BBAW

Redaktion: Stephan Leibfried mit Ute Tintemann

Grafik: angenehme gestaltung/Thorsten Probst

Druck: Druckerei Conrad, Berlin

© Berlin-Brandenburgische Akademie der Wissenschaften, 2013

Jägerstraße 22–23, 10117 Berlin, www.bbaw.de

Nachdruck, auch auszugsweise, nur mit ausdrücklicher Genehmigung des Herausgebers

ISBN: 978-3-939818-36-6

INHALTSVERZEICHNIS

Vorwort	7
1. Einleitung	13
2. Die hegemoniale Stellung der Drittmittel als Indikator zur Messung von Forschungsleistungen	18
2.1 Das DFG-Förderranking	24
2.2 Die Exzellenzinitiative	28
2.3 Leistungsorientierte Mittelvergabe durch die Bundesländer	31
2.4 Berufungspraxis	33
3. Was taugen Drittmittel als Indikator zur Messung von Forschungsleistungen?	35
3.1 Theoretische Kritik: Drittmittel messen den Input und nicht den Output von Forschung	35
3.2 Empirische Kritik: Der statistische Zusammenhang zwischen Drittmitteln und Ergebnismessungen von Forschungen	37
3.3 Kritik an der Art und Weise der Berechnung und Nutzung des Indikators Drittmittel	43
3.4 Ein Sonderweg in der Leistungsmessung ist nicht sinnvoll	46
4. Zusammenfassung und Ausblick	49
Literatur	53

Vorwort

Die vier wichtigsten Bund-Länder-Programme für die Forschung (Hochschulpakt, Exzellenzinitiative, Hochschulbau, Pakt für Forschung und Innovation) enden zwischen 2013 und 2020. Man muss allein mit Blick auf die Wahlen zum Bundestag im September 2013 annehmen, dass schon jetzt grundlegende Weichen dafür gestellt werden, ob und wie diese Programme, insbesondere die eng zusammenhängenden Pakte für Forschung und Innovation (2015) und zur Exzellenzinitiative (2017), fortgesetzt werden, und dass dann im Wesentlichen nur noch Feinarbeit auf dem Programm steht. Außerdem könnten auch zum Gesamtsystem der vier Hochschulpakete Vorentscheidungen fallen, die die Exzellenzinitiative berühren werden, etwa beim Hochschulpakt 2020 oder bei einer Rückkehr zu einer verstärkten Finanzierung des Hochschulbaus durch den Bund (2013–2019).

Um der Diskussion zur Zukunft der deutschen Wissenschaftspolitik in diesen kritischen Jahren ein Forum zu bieten, hat sich die Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) entschlossen, statt eher intern zirkulierender Arbeitspapiere eine eigene Schriftenreihe zur „Wissenschaftspolitik im Dialog“ aufzulegen, die sich an das wissenschaftspolitisch interessierte Publikum in Deutschland richtet.

Die Schriftenreihe „Wissenschaftspolitik im Dialog“ wird von der interdisziplinären Arbeitsgruppe (IAG) *Exzellenzinitiative* (EI) der BBAW betreut. Im Titel der Schriftenreihe wird der Dialog betont, weil in der Reihe ausdrücklich auch Beiträge aus der Wissenschaftspolitik willkommen geheißen werden, in denen bundesweite Vorstellungen zur Zukunft des Wissenschaftssystems entwickelt werden.

Wir freuen uns, mit dem Papier unseres Arbeitsgruppenmitglieds Jürgen Gerhards eine kritische Auseinandersetzung mit der Rolle von Drittmitteln als Indikator zur Messung von Forschungsleistungen vorlegen zu können. Gerhards diskutiert die dominante Rolle, die die Drittmittelinwerbung vor allem auch durch die Förderrankings der DFG bei der Bewertung von Forschungsleistungen in Deutschland spielt. Aus seiner Sicht bildet der Drittmittelindikator nicht das eigentliche Ziel der Forschung, nämlich den Erkenntnisfortschritt ab, der sich eher in der Messung von Publikationen und Zitationen widerpiegelt. Gerhards schlägt vor, in Zukunft verstärkt bibliometrische Verfahren der Leistungsmessung bei der Bewertung von Forschungsleistungen einzubeziehen, um die Qualität

von Forschung, auch mit Blick auf den unterschiedlichen Drittmittelbedarf der unterschiedlichen Fächer, angemessen bewerten zu können.

Seit ihrer Einrichtung im Jahr 2008 hat die IAG die Exzellenzinitiative kritisch begleitet und ihre erste Arbeitsperiode 2010 mit einer Buchveröffentlichung abgeschlossen.¹ Mit der Fortsetzung der Exzellenzinitiative 2011 beschloss die BBAW, die IAG zunächst für drei Jahre fortzuführen, allerdings immer mit dem Jahr 2017 im Visier, in dem die jetzige EI in ihrer bisherigen Struktur ausläuft und in dem Entscheidungen über eine Fortsetzung oder einen neuen Anlauf umzusetzen wären.

Noch im Jahr 2011 gingen wir in der IAG davon aus, wir könnten uns mit unserer kritischen Begleitung der zweiten Exzellenzinitiative Zeit lassen, weil ihre große Evaluation durch die DFG und den Wissenschaftsrat erst im Jahr 2015 ansteht. Ginge alles rational zu, müsste ja jede Entscheidung darüber, was auf die Exzellenzinitiative nach 2017 folgen soll, darauf aufbauen.

Nun sind die Dinge aber schon seit Ende 2011 deutlich immer schneller in Bewegung gekommen:

- So kursiert seit Herbst 2011 der Vorschlag, die Charité-Universitätsmedizin Berlin mit dem Max-Delbrück-Centrum für Molekulare Medizin (MDC) in Berlin-Buch, einem Forschungszentrum der Helmholtz-Gemeinschaft, in einer solchen Form zusammenzuführen, dass eine Einrichtung eigener Art entsteht, die auch mit Bundesmitteln finanziert wird und so die Forschungsfinanzierung des Landes entlastet.² Die entsprechenden laufenden Planungen könnte man als Helmholtz-Eigenlauf *oder* als Vorgriff auf eine „Exzellenzstrategie des funktionalen regionalen Verbunds“ als mögliche Grundausrichtung einer EI 3.0 ab 2017 verstehen. Der eigentliche wissenschaftspolitische Mehrwert des **Charité-„Verbunds“** müsste allerdings darin bestehen, die KIT (Karlsruhe Institute of Technology)-Kombipackung von zwei organisatorisch vereinten, aber sachlich immer getrennten Organisations- und Finanzkreisläufen (des Landes und des Bundes) zu überwinden und zu einer universitär mitgeprägten neuen Verbund-Synthese zu finden.

1 Stephan Leibfried, Hrg., Die Exzellenzinitiative: Zwischenbilanz und Perspektiven, Frankfurt a. M./New York: Campus 2010.

2 www.charite.de/charite/presse/pressemitteilungen/artikel/detail/charite_und_mdc_buendeln_ihre_spitzenforschung_unter_einem_dach/ (Zugang 27.3.2012). Inzwischen ist die Umsetzungsphase erreicht. Eine entsprechende Verwaltungsvereinbarung findet sich im Web in den Drucksachen des Berliner Abgeordnetenhauses (Drs. vom 10.1.2013, 17. Wahlperiode, Nr. 0747).

- Im Januar 2012 fand der Vorsitzende des **Wissenschaftsrats**, Wolfgang Marquardt, unter der Überschrift „**funktionale Verbände in einer Region**“ in einer Rede eine Formel, die eine solche Exzellenzflugbahn als Testballon skizzierte und schon auf die Jahre 2017ff. hin verallgemeinerte: „Viele Regionen in Deutschland sind durch eine reichhaltige Präsenz unterschiedlichster Wissenschaftseinrichtungen – etwa Universitäten, Fachhochschulen oder verschiedenartige außeruniversitäre Forschungseinrichtungen – geprägt. Bei dem Begriff der Region sollten Sie dabei nicht nur an eine Stadt (wie Berlin) oder an einen Ballungsraum (wie das Ruhrgebiet) denken, es kann sich auch – durchaus länderübergreifend – um ein geographisch sinnvoll begrenztes Gebiet mit mehreren Standorten handeln. Nach funktionalen Gesichtspunkten sorgfältig ausgewählte Einrichtungen einer Region könnten, müssten dazu zunächst ihre spezifischen Stärken (z.B. in der Lehre, der Spitzenforschung, der Forschung in kleinen Fächern, im Ergebnistransfer usw.) komplementär ausprägen und dann einen kooperativen Verbund bilden, der sich durch eine breite Funktionalität auf hohem qualitativen Niveau auszeichnet. [...] Offenheit der Einrichtungen, eine intelligente, administrative Reibungsverluste vermeidende Governance und finanzielle Anreize sind Voraussetzungen für eine solche weitreichende Transition und deren Erfolg. Ich kann mir sehr gut vorstellen, dass über eine solche Maßnahme – vielleicht gar als konsequente Weiterführung der Exzellenzinitiative nach 2017 – eine weitere Erhöhung der Leistungsfähigkeit des Wissenschaftssystems in Deutschland erreicht werden kann. Sie würde verschiedenste ‚Exzellenzkriterien‘ bedienen und neben der Spitze auch die ‚imaginäre Mitte‘ im Blick haben.“³
- Im Januar 2012 richtete der **Wissenschaftsrat** eine neue **Arbeitsgruppe** „**Perspektiven der deutschen Wissenschaft**“ ein, die sich aus Anlass des Zusammentreffens des Auslaufens diverser Bund-Länder-Pakte mit dem Ende des Solidarpakts Ost und dem Eintreten der Schuldenbremse der Frage annehmen soll, wie sich das Wissenschaftssystem auf diese massiv veränderten Rahmenbedingungen einstellen kann und soll.⁴

³ Wolfgang Marquardt, Perspektiven der Wissenschaftsfinanzierung in Deutschland nach 2017, Rede: Drs. Köln, 06.01.2012 /Goll / VS, vv. Man., 13 S., hier S. 12f.

⁴ Da sich das in dieser Reihe erschienene Arbeitspapier von Peter Gaehtgens und die Aufgabenstellung der Arbeitsgruppe des Wissenschaftsrats in vielem decken, sei das Mandat der Wissenschaftsratsgruppe genauer zitiert: „Deutschland verfügt über ein umfangreiches und institutionell vielfältiges System wissenschaftlicher Einrichtungen, die von Bund und Ländern in teils besonderer, teils gemeinsamer Verantwortung getragen und finanziert werden. Mit Blick auf die sich stetig wandelnden Aufgaben und Rahmenbedingungen

Diese Arbeitsgruppe tagte erstmals Anfang 2012. Ihr Bericht wird für die Sitzung vom 24.–26. April 2013 erwartet.

- Ferner ergriff die Bundesregierung – vorbereitet durch bayerische und schleswig-holsteinische Initiativen – im März 2012 eine Initiative, Art. 91b I 1 Nr. 2 Grundgesetz dahingehend zu ändern, dem Bund künftig langfristige Ko-Finanzierungsoptionen nicht allein für „Vorhaben“, sondern auch für „Einrichtungen an Hochschulen“ einzuräumen.⁵ Die SPD stimmte dem zwar zu, verlangte aber im Gegenzug, über einen neu einzufügenden Art. 104c Grundgesetz „dauerhafte Finanzhilfen des Bundes für Bildung“ auf Grund von „Vereinbarungen“ zu ermöglichen, die „von den Ländern nur einstimmig beschlossen werden können“ (BT Drs. 17/8455).
- Und nicht zuletzt werfen die Bundestagswahlen im September 2013 ihren langen Schatten in Positionierungen der Parteien und in potentiellen Koalitionslinien zur Wissenschafts- und Hochschulpolitik und damit auch zur Exzellenz-Nachfolge voraus.

All diese Entwicklungen haben zu dem Beschluss geführt, unsere Arbeit zu beschleunigen und so früh wie möglich unsere Arbeitsergebnisse in einer solchen Schriftenreihe öffentlich zugänglich zu machen.⁶

und auf neue Kooperationsformen zwischen wissenschaftlichen Einrichtungen sowie die Diskussionen zur Weiterentwicklung der Wissenschaftsfinanzierung durch Bund und Länder aufbauend auf den gemeinsamen Programmen (bspw. Exzellenzinitiative, Hochschulpakete, Pakt für Forschung und Innovation) ist es deshalb entscheidend, die Funktionalität und Arbeitsteilung der einzelnen Typen wissenschaftlicher Einrichtungen und ihr Zusammenwirken auf strukturelle Schwächen hin zu analysieren, um auf dieser Basis für das gesamte Wissenschaftssystem förderliche Entwicklungsrichtungen aufzeigen zu können. Zu klären gilt es, in welcher Art und welchem Umfang Wissenschaft künftig an Hochschulen oder außeruniversitär, in welchem Verhältnis institutionell oder projektbezogen gefördert wird. Dabei müssen auch die zunehmenden Einflüsse der Europäisierung und Internationalisierung der Wissenschaft auf die wissenschaftlichen Einrichtungen berücksichtigt werden. Vor diesem Hintergrund hat der Wissenschaftsrat im Januar 2012 eine Arbeitsgruppe eingerichtet und beauftragt, längerfristige Perspektiven der wissenschaftlichen Einrichtungen in Deutschland, ihrer Organisation und ihrer Förderung zu entwickeln und zu bewerten. Die Beratung der Empfehlungen im Wissenschaftsrat ist für April 2013 vorgesehen.“ (www.wissenschaftsrat.de/arbeitsbereiche-arbeitsprogramm/forschung/#c6122; Vorsitz: Wolfgang Marquardt, Betreuung in der Geschäftsstelle: Moritz Mälzer; Zugang 27.3.2012).

5 Vgl. u.a. www.bmbf.de/de/17975.php. Wenn man „Vorhaben“ hinreichend weit verstehen will, könnte das „Einrichtungen“ mit umfassen.

6 Zu einem ersten und schnellen Überblick zur heutigen Lage vgl. Stephan Leibfried, Durch die Mitte zur Spitze – Quo vadis 2017, Exzellenzen? Von der dualistischen zur monistischen Innen-Außen-Kooperation in der Forschung, in: Gegenworte 28, Herbst 2012 (Schwerpunktheft „Zwischen den Wissenschaften“), S. 31–35. Zum aktuellen Stand s. Heike Schmolli, Hungrige Universitäten, Frankfurter Allgemeine Zeitung vom 18. März 2013 (Nr. 65), S. 4.

Vielleicht sollten wir heute so schließen wie schon 2010: „Die Ivy League entstand nicht in zwei mal fünf Jahren, sondern durch langsame Institutionenbildung und lange Investitionsketten. Die Exzellenzinitiative kann nur ein Anfang sein. Wo, wenn nicht in der Wissenschaftspolitik, wäre ‚das langsame Bohren dicker Bretter mit Leidenschaft und Augenmaß‘ (Max Weber) die einzig erfolgversprechende Handlungsorientierung?“⁷

Günter Stock

Präsident der Berlin-Brandenburgischen Akademie der Wissenschaften

Stephan Leibfried

Sprecher der interdisziplinären Arbeitsgruppe Exzellenzinitiative der BBAW

⁷ Stephan Leibfried und Günter Stock, Vorwort, in: Die Exzellenzinitiative (siehe Fußnote 1), S. 7–9, hier S. 9.

DER DEUTSCHE SONDERWEG IN DER MESSUNG VON FORSCHUNGSLEISTUNGEN¹

1. EINLEITUNG²

Eine Bewertung der Forschungsqualität einzelner Wissenschaftler, aber auch von ganzen Instituten und Universitäten hat es aus historischer Sicht immer schon gegeben. Der Ruf prominenter Wissenschaftler im jeweiligen Forschungsfeld hallte diesen voraus und zog Studierende, Doktoranden und Kollegen an, führte zur Etablierung von Forschungszentren und damit häufig zu einer Hierarchisierung des Systems. Seit den 1990er Jahren wurden aber zunehmend *systematische* Verfahren der Dauerbeobachtung und Bewertung von wissenschaftlichen Leistungen entwickelt und institutionalisiert (Gläser 2008; Gläser u. a. 2008; Hicks 2012).

Zu diesem Prozess der Institutionalisierung der Dauerbeobachtung von Wissenschaft gehört erstens die Definition von Indikatoren, mit denen versucht wird, wissenschaftliche Qualität zu einer messbaren und vergleichbaren Größe zu machen. So sind Zitationsmessungen und das Ausmaß eingeworbener Drittmittel zentrale Bestandteile der entwickelten Maßzahlen. Die definierten Verfahren und Indikatoren wurden zweitens flächendeckend auf die Bewertung kleinerer und größerer Einheiten angewandt. Dabei reicht die Spannweite der erfassten und verglichenen Einheiten von einem internationalen Vergleich von Universitäten über einen Vergleich der nationalen Forschungseinrichtungen bis hin zu binnenuiversitären Vergleichen einzelner Institute. So vergleicht z. B. das umstrittene „Academic Ranking of World Universities“, das sogenannte Shanghai Ranking³ und das „Times Higher Education World University Ranking“⁴

1 Der Begriff des deutschen Sonderwegs kommt aus der Geschichtswissenschaft und ist zur Beschreibung der deutschen Entwicklung im 19. und 20. Jahrhundert benutzt worden (vgl. zusammenfassend Kocka 1988). Die Sonderwegsthese ist in der Geschichtswissenschaft überaus umstritten. Ich benutze den Begriff hier allerdings ausschließlich metaphorisch, um zum Ausdruck zu bringen, dass in Deutschland anders als in anderen Ländern den eingeworbenen Drittmitteln eine übergebührende Bedeutung bei der Messung von Forschungsqualität zugesprochen wird.

2 Ich bedanke mich bei den Mitgliedern der Interdisziplinären Arbeitsgruppe *Exzellenzinitiative* der Berlin-Brandenburgischen Akademie der Wissenschaften für hilfreiche Kommentare, besonders aber bei Friedhelm Neidhardt, Jochen Gläser und Stephan Leibfried für eine schriftliche Kommentierung einer Erstfassung des Manuskripts. Dank geht auch an Inga Ganzer, die mich bei der Datenrecherche unterstützt hat.

3 www.shanghairanking.com/ARWU2012.html

4 www.timeshighereducation.co.uk/world-university-rankings

viele Universitäten weltweit miteinander. Das von der Deutschen Forschungsgemeinschaft publizierte Förderranking⁵ beschränkt sich auf den Vergleich deutscher Universitäten und das vom Wissenschaftsrat verantwortete Rating bezieht sich auf den Vergleich einzelner Fächer innerhalb der Bundesrepublik.⁶

Eine systematische Evaluation von Forschungsleistungen bedeutet drittens, dass die meisten Erhebungen Längsschnittstudien sind und in regelmäßigen Abständen wiederholt werden. Dadurch ist ein Vergleich über die Zeit möglich, so dass sich in der Spitzengruppe Aufsteiger und Absteiger identifizieren lassen. Schließlich und viertens geht der Prozess der systematischen Evaluation von Forschungsleistungen mit der Entstehung und Institutionalisierung einer Evaluationsindustrie einher, der Gründung einer Vielzahl von Organisationen mit zum Teil großen Mitarbeiterstäben, die als unabhängige Unternehmen agieren oder als neue Abteilungen innerhalb bestehender wissenschaftlicher Institutionen, wie die der DFG, des Wissenschaftsrats und der Universitäten, gegründet wurden.

Die systematische, indikatorenbasierte Dauerbeobachtung und Bewertung wissenschaftlicher Leistungen ist selbst wiederum in einen größeren gesellschaftlichen Entwicklungstrend eingebettet. Auch in anderen Bereichen haben sich, zum Teil im Weltmaßstab, Institutionen der Dauerbeobachtung und Bewertung von teilsystemspezifischen Leistungen entwickelt, die mit der Veröffentlichung ihrer Bewertungen auf die Entwicklung der Gesellschaft einwirken. Die Kreditwürdigkeit von Banken und ganzer Länder wird durch Rating-Agenturen beobachtet (Botzem 2012; Krück 2011), die Leistungen von Schülern werden ländervergleichend durch einschlägige Bildungsstudien wie PISA beurteilt. Wohlfahrtsstaaten, Demokratien und Gesundheitssysteme werden systematisch anhand von Performanzindikatoren verglichen; und über das Niveau von Korruption in den Ländern der Welt gibt der Korruptionsindex von „Transparency International“ Auskunft. Die beständige Bewertung von Leistungen geht hinunter bis auf die Ebene der Evaluation einzelner Personen, so dass man z. B. die Bewertung von Ärzten und deren Qualität und die von Professoren und ihrer Lehre im Internet nachvollziehen kann. Michael Power (1997) hat zur Beschreibung dieses allgemeinen Entwicklungstrends den Begriff der „Audit-Society“ eingeführt.

5 www.dfg.de/dfg_profil/foerderatlas_evaluation_statistik/foerderatlas/index.jsp

6 www.wissenschaftsrat.de/arbeitsbereiche-arbeitsprogramm/forschungsrating.html

Die Etablierung von Verfahren der flächendeckenden Dauerbeobachtung wissenschaftlicher Leistungen und die Publikation ihrer Ergebnisse hat das Wissenschaftssystem nachhaltig verändert. Allein die Zuweisung einer hohen Rangposition, sei es im internationalen, im nationalen oder auch im universitätsinternen Vergleich, ist eine Auszeichnung, die als Qualitätsmerkmal für die jeweilige Forschungsleistung von den Betroffenen und dem Umfeld interpretiert wird. Der Rang führt zur Vermehrung des symbolischen Kapitals desjenigen, der ausgezeichnet wird, und ist mit einer Reputationssteigerung verbunden. Diese Reputationssteigerung kann wiederum häufig in andere „Währungen“ konvertiert werden. Ausgezeichnete Institutionen werden für Studierende, Doktoranden und Kollegen attraktiver, was dazu führt, dass bessere Leute rekrutiert werden, was sich längerfristig wiederum in einer Verbesserung der Forschungsleistungen und dann auch der Evaluationsergebnisse auswirken wird. Es erfolgt ein Prozess der positiven Selbstverstärkung, der bekanntlich von Robert K. Merton (1968) als Matthäus-Effekt bezeichnet wurde.

Eine hohe Rangposition ist häufig auch mit materiellen Gratifikationen verbunden und zwar sowohl für die Institutionen als auch für die Individuen. So koppelnd zunehmend mehr Länder die Finanzierung von Universitäten und außeruniversitären Forschungseinrichtungen an das Abschneiden in Evaluationsverfahren. Institute und Universitäten, die als erfolgreich ausgezeichnet worden sind, bekommen mehr Ressourcen und können damit wiederum ihre Lehre und ihre Forschung verbessern. Effekte zeigen sich aber auch auf der Individualebene. Die durch Evaluationsverfahren ausgezeichneten Wissenschaftler können höhere Gehälter oder Leistungszulagen fordern; sie verbessern zudem ihre Chancen, an andere und häufig höher klassifizierte Universitäten berufen zu werden, die wiederum meist auch bessere Gehälter zahlen.

Insofern handelt es sich bei der Institutionalisierung der Dauerbeobachtung und Bewertung von Forschungsleistungen nicht um ein folgenloses Glasperlenspiel, sondern um ein Belohnungssystem, das verhaltenssteuernd wirkt und langfristig zu Strukturveränderungen führt. Nicht alle durch die Dauerbeobachtung der Wissenschaft ausgelösten Veränderungen sind dabei erfreulich. Einen guten Überblick über Strukturveränderungen mit einem besonderen Augenmerk auf die dysfunktionalen Effekte findet sich in einem von der OECD (2010) veröffentlichten Bericht (vgl. auch Gläser u. a. 2008; Hicks 2012).

Dass sich Universitäten, Fakultäten und die einzelnen Fächer einer systematischen Überprüfung ihrer Qualität stellen müssen, ist im Grundsatz eine gute Sache. Die schrittweise und flächendeckende Implementierung von Evaluationsverfahren hat in Deutschland sicherlich dazu beigetragen, den Dornröschenschlaf der deutschen Universitäten zu beenden. Gerade weil der Leistungsmessung aber eine so wichtige Rolle zukommt, ist es entscheidend, welche Indikatoren benutzt werden, um die Forschungsleistungen von Universitäten und außeruniversitären Einrichtungen zu messen. Indikatoren sind Anzeichen von etwas, sie sind nicht das Phänomen selbst, das sie anzeigen und zu messen vorgeben. Entsprechend bedeutsam ist es, ob die zur Messung der Leistungen von Universitäten, Instituten und auch einzelner Personen benutzten Indikatoren valide Messungen ihrer faktischen Leistungen sind.

Ich werde im zweiten Abschnitt an vier verschiedenen Beispielen skizzieren, dass Deutschland in der Bemessung der Forschungsqualität einen besonderen Weg eingeschlagen hat, weil hier vor allem die eingeworbenen Drittmittel als zentraler Indikator zur Messung von Forschungsleistungen herangezogen werden. In vielen anderen Ländern ist dies nicht bzw. nicht in diesem Maße der Fall. Dort spielen Publikationen und deren Zitationen eine wesentlich wichtigere Rolle bei der Messung von Forschungsleistungen.

Ob und in welchem Maße die eingeworbenen Drittmittel eine valide Messung von Forschungsleistungen darstellen, wird im dritten und zentralen Abschnitt diskutiert. Vier Argumente stehen hier im Vordergrund: **(1)** Drittmittel sind ein Produktionsfaktor der *Ermöglichung* von Forschung, sie messen aber nicht deren Ergebnis und Qualität. Insofern ist es theoretisch wenig plausibel, Drittmittel als Indikator zur Messung von Forschungsergebnissen zu benutzen. **(2)** Obwohl wir erstaunlich wenig über den korrelativen Zusammenhang zwischen Drittmitteln und anderen Messungen von Forschungsleistungen wissen, zeigen die wenigen Studien, die vorliegen, dass dieser Zusammenhang eher gering ist. Insofern spricht auch empirisch wenig dafür, Drittmittel als Ersatzmessung von Forschungsleistungen zu benutzen. **(3)** Hinzu kommt, dass auch die spezifische Berechnung der Drittmittelquote, wie sie vielerorts üblich ist, Probleme aufweist. Weder scheint es sinnvoll zu sein, Drittmittel als absolute Größe auszuweisen ohne sie ins Verhältnis zum wissenschaftlichen Personal zu setzen, noch ist es überzeugend, die eingeworbenen Drittmittelsummen zu berechnen; besser wäre es, nicht die Geldbeträge, sondern die Anzahl der eingeworbenen Projekte auszuweisen. **(4)** Schließlich werde ich argumentieren, dass durch die Tatsache, dass

Drittmittel besonders in Deutschland als zentrale Größe zur Leistungsmessung definiert werden, die falschen Anreize gesetzt werden, so dass der Leistungsabstand im Vergleich zu anderen Ländern sich in Zukunft eher vergrößern als verkleinern könnte.

In einem vierten Abschnitt fasse ich die Ergebnisse zusammen und gebe einen kurzen Ausblick.

Um möglichen Missverständnissen und Einwänden vorzubeugen, sind vorweg einige Spezifikationen und Erläuterungen notwendig. **(a)** Die hier formulierte Kritik an der deutschen Drittmittelfixiertheit bezieht sich auf die Nutzung des Drittmittelaufkommens als alleinigem bzw. dominantem Indikator zur Messung der Qualität von Forschung. Mit der Kritik ist nicht gemeint, dass Drittmittel für bestimmte Fächer und in Kombination mit publikationsbasierten Indikatoren nicht auch zur Messung der Forschungsqualität genutzt werden können. **(b)** Diese Kritik impliziert auch nicht, dass Drittmittelindikatoren für andere Zwecke keine sinnvolle Funktion erfüllen und eine valide Messgröße darstellen können. Der Wissenschaftsrat (2008) unterscheidet in seinem Forschungsrating verschiedener Fächer u. a. zwei Dimensionen der Bewertung: „Forschungsqualität“ einerseits und „Transfer von Forschung in andere gesellschaftliche Bereiche“ andererseits. Die hier vorgetragene Kritik bezieht sich auf die Nutzung von Drittmitteln als Indikator zur Messung der Forschungsqualität und nicht auf die Nutzung des Indikators zur Messung von Transfer. Bei einem Teil der Forschung handelt es sich bekanntlich um anwendungsorientierte Forschung; diese zielt auf die Lösung von konkreten, praktischen Problemen und fällt unter die Kategorie „Transfer“. **(c)** Im Jahr 2009 belief sich das Drittmittelaufkommen der deutschen Universitäten auf 5,3 Mrd. Euro, während die Grundmittel bei 20 Mrd. lagen (DFG 2012: 29). Ohne Drittmittel könnte ein großer Teil der Forschung gar nicht stattfinden, der wissenschaftliche Arbeitsmarkt würde zusammenbrechen. Die zunehmende Bedeutung der Drittmittel für die Ausstattung der Universitäten soll mit der hier vorgetragenen Kritik nicht in Abrede gestellt werden.⁷ Es geht allein um die Frage, ob Drittmittel ein guter Indikator zur Messung der Forschungsqualität sind.

⁷ Während sich die von Hochschulen eingeworbenen Drittmittel zwischen 1998 und 2009 um mehr als 100% erhöht haben, sind im selben Zeitraum die Grundmittel um 23% angestiegen (DFG 2012: 29).

Die hier angestellten Überlegungen beanspruchen nicht für sich, überaus originell zu sein. Sie stützen sich auf Diskussionen und Forschungen von Kolleginnen und Kollegen (z. B. Gläser 2008; Gläser u. a. 2004, 2008; Hicks 2012; Hornbostel 1997, 2001; Jansen u. a. 2007; Laudel 2006; Moed u. a. 1998; Menninghaus 2006; Münch 2006, 2007; Neidhardt 2006, 2010, 2012), tragen diese zusammen und bündeln sie zu einer These mit einer wissenschaftspolitischen Botschaft: Die hegemoniale Rolle, die Drittmittel in Deutschland bei der Bemessung der Forschungsqualität spielen, ist nicht sinnvoll und setzt die falschen Anreize. Dabei haben die Zentralakteure des deutschen Wissenschaftssystems, vor allem die Deutsche Forschungsgemeinschaft, aber auch der Wissenschaftsrat, einen entscheidenden Anteil daran, dass Drittmittel als Forschungsleistungsindikator fetischisiert⁸ werden. Die Entwicklung alternativer Indikatoren, die sich an der Qualität von Publikationen orientiert, ist dringend geboten.

2. DIE HEGEMONIALE STELLUNG DER DRITTMITTEL ALS INDIKATOR ZUR MESSUNG VON FORSCHUNGSLEISTUNGEN

Ein zentrales Ziel wissenschaftlicher Forschung ist die systematische Produktion von Erkenntnissen über unterschiedliche Gegenstandsbereiche und damit die Verbesserung des Wissens über die Welt. Innerhalb der Wissenschaftsforschung gibt es weitgehend Konsens darüber, dass das, was eine Verbesserung des Wissens genau ist, objektiv nicht feststellbar ist. Gleichsam als Ersatzmessung von objektiver Erkenntnis und Erkenntnisfortschritt wurde die Idee des „peer review“ institutionalisiert: Mehrere unabhängige Gutachter, die selbst Experten in dem jeweiligen Forschungsfeld sind, kommentieren und bewerten die Forschungsleistungen anderer. Sowohl die Gutachter als auch die Begutachteten bleiben meist anonym. Das Urteil der Experten ersetzt damit gleichsam das Kriterium der Wahrheit (vgl. Weingart 2001: 285; Luhmann 1972; Hornbostel 1997; Neidhardt 2010).

Die Ergebnisse von wissenschaftlichen Forschungen werden in aller Regel publiziert oder materialisieren sich in neuen technischen Produkten und Erfindungen. Das Urteil der Fachkollegen bei der Bemessung der Qualität von Forschungen bezieht sich entsprechend auf das Forschungsergebnis, so wie es in Publikationen und Erfindungen zum Ausdruck kommt. Die Qualität der durch

⁸ Der Begriff des Fetischs zur Beschreibung der Drittmittelorientierung ist jüngst auch von Stefan Kühl (2013) in einem Artikel in der Süddeutschen Zeitung benutzt worden.

Fachkollegen zertifizierten Forschungsergebnisse kommt dabei in zweifacher Weise zum Ausdruck:

- a) Die Veröffentlichung des über „Peer-Review“-Verfahren zertifizierten Wissens erfolgt über Bücher, vor allem aber über Fachzeitschriften. Zum Ende des 18. Jahrhunderts entstanden, haben sich heute für fast alle wissenschaftlichen Disziplinen verschiedene Fachzeitschriften etabliert, die untereinander nach Reputation – und meist damit zusammenhängend: nach Ablehnungsquote eingereichter Manuskripte – hierarchisiert sind (Stichweh 1988). Artikel in Fachzeitschriften sind vor allem für die Naturwissenschaften das zentrale Forschungsprodukt. Aber auch in den Geistes- und Sozialwissenschaften gewinnen Zeitschriften in Relation zu der für diese Wissenschaften typischen Publikationsform – der Monographie – an Bedeutung. Der Zugang zu den Fachzeitschriften wird durch ein „Peer-Review“-Verfahren überwacht. Mehrere Gutachter äußern sich zum Teil in sehr ausführlicher und detaillierter Weise über die bei den Zeitschriften eingereichten Artikel und entscheiden mit über die Publikationswürdigkeit einer wissenschaftlichen Abhandlung. Insofern ist die Anzahl an Publikationen in Fachzeitschriften, die zu den sogenannten „Peer-Review“-Journals gehören, ein wichtiger Indikator der Qualitätsmessung wissenschaftlicher Forschungsleistung, weil einerseits der Bewertungsprozess am Endprodukt der Forschung (Publikationen) ansetzt, andererseits eine Qualitätskontrolle durch Fachkollegen stattfindet.

- b) Liegen erst einmal Veröffentlichungen vor, die durch den Prozess der Qualitätskontrolle gegangen sind, dann ist entscheidend, ob die Veröffentlichungen auch die Forschungen anderer beeinflussen, ob also deren Erkenntnisse von anderen Personen, die im Feld arbeiten, überhaupt wahrgenommen und aufgenommen werden. Die wissenschaftliche Wahrnehmung der Leistungen anderer kommt in Zitationen zum Ausdruck, in der Erwähnung der Forschungen von Kolleginnen und Kollegen in den eigenen Publikationen. Entsprechend sind Zitationen von Veröffentlichungen anderer Wissenschaftler ein weiteres Merkmal der Qualitätsmessung.

Die Grundidee, Forschungsleistungen durch das Urteil von Experten, die diese im Hinblick auf den „Outcome“ von Forschungen abgeben, zu bemessen, hat ohne Zweifel eine hohe Plausibilität und auch eine gewisse „Ableitungseleganz“. In der Sprache der Theorie sozialer Produktionsfunktionen kann man das

Argument folgendermaßen reformulieren (Lindenberg 1996; Esser 1999): Das eigentliche Produkt des Wissenschaftssystems besteht in der Herstellung von Erkenntnisfortschritt. Dieser ist objektiv nicht messbar. Entsprechend bedarf es der Definition von primären Zwischengütern. Dazu gehören Publikationen und die Beurteilung der Qualität von Publikationen durch entsprechende Experten im Feld. Es spricht erstmal nichts dafür, sekundäre oder tertiäre Zwischengüter – wie z. B. die Einwerbung von Drittmitteln – als Indikatoren zur Messung von Forschungsleistungen einzuführen. Diese sind in der Kette der Zwischengüter gleichsam weiter hinten gelagert; und es ist unwahrscheinlicher, dass sie die Qualität von Forschungen besser messen als die primären Zwischengüter. Dass auch mit der Messung von Publikationen und Zitationsindizes eine Vielzahl an Problemen und Ungenauigkeiten verbunden ist, soll hier nicht bestritten werden. In Relation zu drittmittelbasierten Messungen scheint eine – zumindest zusätzliche – Orientierung an Publikationen aber sowohl theoretisch als auch empirisch weitaus plausibler zu sein.

Die meisten international vergleichenden Evaluationsverfahren beziehen sich entsprechend in erster Linie auf Publikationen als zentraler Einheit zur Messung von Forschungsleistungen. Und sie orientieren sich an der Resonanz von Publikationen, wie sie sich in Zitationen manifestieren. Dazu einige Beispiele:⁹ Das sogenannte Shanghai-Ranking (Academic Ranking of World Universities, ARWU), in dem seit 2003 ganze Universitäten bewertet werden, berücksichtigt neben der Menge der Nobelpreise (und für die Mathematik die Anzahl der gewonnenen Fields-Medaillen) vor allem die Zitations- und Publikationshäufigkeiten der Mitglieder einer Universität. Das „Times Higher Education World University Ranking“ publiziert seit 2004 eine Rangfolge von Universitäten. Seit 2010 gibt es eine Kooperation mit Thomson Reuters und seitdem hat sich auch die Methodik geändert. Insgesamt werden 13 Indikatoren zur Leistungsbemessung berücksichtigt. Einer Reputationsbefragung unter Akademikern und einer Auswertung der Zitationshäufigkeit von Publikationen kommt dabei das größte Gewicht zu. Das methodisch elaborierteste international vergleichende Ranking-Verfahren ist das sogenannte „Leiden-Ranking“. Dabei handelt es sich um ein rein bibliometrisches Verfahren. Da die Zitationsraten je nach Disziplin und Subdisziplin enorm variieren, werden hier fachspezifisch normalisierte Zitationsraten berücksichtigt.

⁹ Die Nennung der verschiedenen Rankings bedeutet nicht, dass die methodische Qualität als ausreichend betrachtet wird.

Aber auch auf der Ebene nationaler Evaluationen universitärer Forschungsleistungen spielt die Publikationsorientierung eine sehr wichtige Rolle. Vorreiter in der Einführung von sogenannten „performance based research fundings systems“ war das Vereinigte Königreich (Gläser 2008). Die Performanz der Universitäten wird durch das „Research Exercise Assessment“ evaluiert, das zum ersten Mal 1986 unter der Regierung von Margret Thatcher institutionalisiert wurde. Die Bewertungen der verschiedenen Institutionen werden durch Gutachtergruppen vollzogen. Grundlage der Urteile der Gutachtergruppen bilden Informationen, die sie von den jeweiligen Institutionen erhalten haben. Die wichtigste Informationsquelle sind Publikationen; Drittmitteln kommt keine Bedeutung zu. Für die letzte Bewertung aus dem Jahr 2008 konnte man für jedes Vollmitglied einer Institution vier „research outputs“, meist also Artikel, die zwischen 2001 und 2007 erschienen sind, einreichen. Auf der Grundlage der Evaluationen erfolgt die Mittelzuweisung für die universitäre Forschung.

Ähnliche performanzbasierte Systeme findet man in anderen Ländern. Der Artikel von Diana Hicks (2012) gibt einen guten Überblick über 13 Länder (Australien, Dänemark, Finnland, Norwegen, Belgien, Polen, Slowakei, Schweden, Vereinigtes Königreich, Italien, Portugal, Spanien und Neuseeland) und die Metropole Hongkong. In den einzelnen Ländern werden unterschiedliche Indikatoren benutzt und auch unterschiedlich miteinander kombiniert. Publikationen und deren Zitationen spielen aber in allen Ländern eine bedeutsame Rolle. Allein Australien scheint dasjenige Land zu sein, in dem ähnlich wie in Deutschland den eingeworbenen Drittmitteln ein besonderer Status zukommt.¹⁰

Werfen wir schließlich noch einen Blick in die USA, dessen Universitätssystem bzw. dessen Eliteuniversitäten in vielen akademischen Belangen als großes Vorbild auch für die Bundesrepublik dienen. Hier gibt es sehr verschiedene Rankings von Universitäten. Die meisten von ihnen beziehen sich auf die Bewertung der Qualität der Lehrprogramme für „Undergraduates“ und nicht unmittelbar auf die Bewertung der Forschungsqualität. Ein gutes Abschneiden in diesen Rankings ist gerade für die privaten Universitäten von zentraler Bedeutung, da die Ergebnisse einen großen Einfluss auf die Wahlentscheidung der (zahlenden) Studierenden haben. Eines der bekanntesten und einflussreichsten Rankings ist das „U.S. News

10 Die leistungsorientierte Forschungsfinanzierung wurde dort 1995 eingeführt (vgl. Gläser u. a. 2008). Bei der Vergabe der Gelder werden folgende Indikatoren berücksichtigt: 54,8% der Gelder werden nach der Summe der eingeworbenen Drittmittel vergeben, 29,1% nach der Zahl der Studienabschlüsse von Forschungsstudierenden, 8,4% nach der Zahl der Publikationen und 7,7% nach der Zahl der immatrikulierten Forschungsstudierenden.

& World Report College and University Ranking“, das seit 1984 jährlich erscheint.¹¹ Zu seiner Erstellung werden verschiedene Indikatoren benutzt. Der wichtigste Faktor geht auf eine Umfrage unter Universitätsmitgliedern zurück, die die Reputation der verschiedenen Universitäten einschätzen.

Auch zur Messung der Forschungsqualität von Universitäten gibt es in den USA verschiedene Rankings. In allen spielen die eingeworbenen Drittmittel aber nicht die zentrale Rolle wie in der Bundesrepublik. Dies sei an einem Beispiel illustriert. Das „United States National Research Council Ranking“ gilt auf Grund seiner methodischen Elaboriertheit und seiner akademischen Akzeptanz gleichsam als der Goldstandard.¹² Es versucht, die Qualität von 5.000 Doktoranden-Programmen in 62 Fächern, angesiedelt an 212 Universitäten, zu bewerten. Für das Ranking werden insgesamt 21 Indikatoren benutzt; diese wird drei Dimensionen zugeordnet: Qualität der Mitglieder des Departments, Merkmale der Studierenden und Merkmale des Doktorandenprogramms. Da wir in unserem Kontext vor allem an der Messung von Forschungsleistungen interessiert sind, konzentriere ich mich im Folgenden allein auf die Indikatoren, die zur Messung der Qualität der Forschung der Institute benutzt werden. Dazu werden wiederum sieben verschiedene Indikatoren herangezogen, die mit einer sehr unterschiedlichen Gewichtung in die Analyse eingehen. Um die Gewichtung zu bestimmen, wurde eine Umfrage unter den Professoren der Fächer durchgeführt, die die Bedeutsamkeit der Indikatoren bestimmen konnten. Auf der Grundlage der Einschätzung der Fachkollegen erhalten die Indikatoren ein spezifisches Gewicht, das je nach Fach auch unterschiedlich ausfällt.

In der folgenden Tabelle habe ich für drei verschiedene Fächer (Physik, Geschichte, Soziologie) rekonstruiert, mit welchem prozentualen Gewicht die einzelnen Indikatoren in die Messung der Qualität der Institute eingehen.

11 <http://colleges.usnews.rankingsandreviews.com/best-colleges>

12 <http://sites.nationalacademies.org/PGA/Resdoc>

Tabelle 1: Die Messung der Forschungsqualität in den USA am Beispiel des „National Research Council Assessment of Doctorate Programs“ (eigene Berechnungen)

	Physics	History	Sociology
Publications per Allocated Faculty	22,65%	36,19%	29,01%
Cites per Publication	26,28%	–	19,49%
Grants per Allocated Faculty	26,90%	13,76%	18,25%
Percent Faculty Interdisciplinary	4,16%	8,97%	6,44%
Percent Non-Asian Minority Faculty	0,97%	8,15%	6,70%
Percent Female Faculty	1,95%	7,85%	5,11%
Percent Faculty Interdisciplinary	17,08%	25,08%	14,99%

Ich möchte hier die Indikatoren nicht im Einzelnen diskutieren, sondern konzentriere mich auf die Bedeutung der Drittmittel in Relation zu den Publikationen und Zitationen. Trotz der Unterschiede zwischen den drei Fächern sieht man, dass den Drittmitteln bei Weitem nicht die Bedeutung zukommt wie den beiden Maßzahlen, die publikationsorientiert sind, nämlich die Anzahl an Publikationen und die Zitationen pro Publikation.

Ich benutze die kurze Beschreibung von Rankings anderer Länder allein als Kontrastfolie, um zu zeigen, dass Deutschland in der Bewertung von Forschungsleistungen einen Sonderweg eingeschlagen hat. Im Vergleich zu den beschriebenen international vergleichenden Evaluationsverfahren, aber auch im Vergleich zu vielen evaluationsbasierten Forschungsrankings, die in anderen Ländern eingeführt wurden, spielen in Deutschland die eingeworbenen Drittmittel die zentrale Rolle in der Beurteilung der universitären Leistungen. Diese besondere Bedeutung soll an vier Beispielen illustriert werden.

2.1 Das DFG-Förderranking

Die Deutsche Forschungsgemeinschaft ist *die* zentrale Selbstverwaltungsorganisation der Wissenschaft in Deutschland. Ihre hegemoniale Stellung verdankt sie nicht allein der materialen Ausstattung mit einem Jahresetat von ca. 2,5 Mrd. Euro, sondern ihrer hohen Reputation. Entsprechend sind die Stellungnahmen der DFG im akademischen Feld mit einer ganz besonderen Dignität und Legitimität ausgestattet und werden in aller Regel von den anderen Institutionen im wissenschaftlichen Feld auch akzeptiert. Seit 1997 veröffentlicht die DFG alle drei Jahre das sogenannte Förderranking (1997, 2000, 2003, 2006, 2009, 2012). Seit 2012 wird es unter dem neuen Namen „Förderatlas – 2012. Kennzahlen zur öffentlich finanzierten Forschung in Deutschland“ publiziert. Mit dem Bericht gibt die DFG einen sehr breiten und detaillierten Überblick über die öffentlich geförderte Forschung für Hochschulen und außeruniversitäre Forschungseinrichtungen. Eingang in den Bericht finden vor allem die Drittmittel-Förderungen der DFG, dann die des Bundes und der EU sowie Personenförderungen der Alexander-von-Humboldt-Stiftung, des Deutschen Akademischen Austauschdienstes und des European Research Council. Zentraler Bestandteil des Berichts ist eine Rangliste von Hochschulen, abgestuft nach den Bewilligungssummen.

Tabelle 2 gibt eine solche Darstellung aus dem DFG-Bericht (2012: 76) wieder. Die Ranglisten werden durch eine Reihe von Differenzierungen weiter aufgeschlüsselt. So wird in der hier wiedergegebenen Tabelle danach unterschieden, ob die Gelder, die Universitäten in der Exzellenzinitiative eingeworben haben, berücksichtigt werden oder nicht. Neben einem Gesamtüberblick, der allein die absoluten Fördersummen pro Universität enthält, findet man in dem Bericht aber auch (um pro Hochschullehrer oder wissenschaftlicher Mitarbeiter) gewichtete Tabellen, eine Differenzierung nach Fachgebieten, Tabellen, die die Rangposition der Universitäten im Zeitverlauf wiedergeben usw. In dem Bericht aus dem Jahr 2012 findet sich auf zehn Seiten eine bibliometrische Analyse für das Fach Chemie, die für 40 Universitäten die Anzahl der eingeworbenen Drittmittel mit der Anzahl an Publikationen in Beziehung setzt.

Tabelle 2: Rangplatzveränderungen bei DFG-Bewilligungen für 2008 bis 2010 mit und ohne Berücksichtigung der Exzellenzinitiative des Bundes und der Länder

RANKINGBETRACHTUNG DER HOCHSCHULEN

Hochschulen mit ExIn				Hochschulen ohne ExIn			
	Rang	Mio. €	kum %		Rang	Mio. €	kum %
Aachen TH	1	278,1	4,7	Berlin FU	1	168,3	3,6
München LMU	2	263,7	9,1	München LMU	2	166,1	7,1
Berlin FU	3	250,8	13,4	Aachen TH	3	159,6	10,5
München TU	4	250,1	17,6	München TU	4	158,6	13,8
Heidelberg U	5	246,3	21,7	Berlin HU	5	149,0	17,0
Freiburg U	6	202,1	25,2	Heidelberg U	6	144,5	20,0
Karlsruhe KIT	7	200,5	28,5	Freiburg U	7	137,3	23,0
Berlin HU	8	179,8	31,6	Dresden TU	8	135,4	25,8
Göttingen U	9	168,6	34,4	Karlsruhe KIT	9	133,1	28,7
Erlangen-Nürnberg U	10	159,1	37,1	Erlangen-Nürnberg U	10	130,7	31,4
Bonn U	11	157,4	39,8	Würzburg U	11	129,1	34,2
Münster U	12	149,4	42,3	Bonn U	12	128,4	36,9
Dresden TU	13	144,0	44,7	Göttingen U	13	127,6	39,6
Tübingen U	14	140,9	47,1	Münster U	14	123,3	42,2
Würzburg U	15	133,1	49,3	Tübingen U	15	119,2	44,7
Frankfurt/Main U	16	131,7	51,5	Köln U	16	107,2	47,0
Köln U	17	129,6	53,7	Bremen U	17	105,5	49,2
Darmstadt TU	18	126,2	55,9	Bochum U	18	100,3	51,4
Konstanz U	19	123,6	57,9	Darmstadt TU	19	99,8	53,5
Hamburg U	20	117,6	59,9	Hamburg U	20	97,4	55,5
Bremen U	21	116,8	61,9	Mainz U	21	93,2	57,5
Hannover U	22	112,4	63,8	Hannover U	22	92,6	59,5
Stuttgart U	23	110,8	65,7	Frankfurt/Main U	23	88,8	61,4
Bochum U	24	104,8	67,4	Stuttgart U	24	86,1	63,2
Mainz U	25	95,9	69,0	Berlin TU	25	80,2	64,9
Berlin TU	26	95,3	70,7	Düsseldorf U	26	77,7	66,5
Kiel U	27	94,4	72,2	Jena U	27	74,8	68,1
Hannover MedH	28	88,5	73,7	Regensburg U	28	72,4	69,6
Bielefeld U	29	78,4	75,1	Duisburg-Essen U	29	70,6	71,1
Düsseldorf U	30	77,7	76,4	Kiel U	30	67,7	72,6
Jena U	31	76,8	77,7	Ulm U	31	67,7	74,0
Regensburg U	32	72,4	78,9	Hannover MedH	32	67,3	75,4
Ulm U	33	70,9	80,1	Marburg U	33	66,9	76,9
Duisburg-Essen U	34	70,6	81,3	Dortmund TU	34	64,5	78,2
Gießen U	35	68,3	82,4	Braunschweig TU	35	61,5	79,5
Marburg U	36	66,9	83,6	Leipzig U	36	59,7	80,8
Dortmund TU	37	64,5	84,6	Gießen U	37	54,8	81,9
Leipzig U	38	63,0	85,7	Bielefeld U	38	53,8	83,1
Saarbrücken U	39	62,5	86,8	Konstanz U	39	53,4	84,2
Braunschweig TU	40	61,5	87,8	Halle Wittenberg U	40	51,9	85,3
Rang 1 – 40		5.204,9	87,8	Rang 1 – 40		4.026,1	85,3
Weitere HS		723,0	12,2	Weitere HS		692,7	14,7
HS insgesamt	186	5.927,9	100,0	HS insgesamt	186	4.718,9	100,0

Trotz dieser Differenzierungen steht aber vor allem ein Indikator im Mittelpunkt der Betrachtung: nämlich der Anteil der eingeworbenen Drittmittel. Die Qualität des Indikators zur Bemessung von Forschungsleistungen ist in der Vergangenheit von der DFG selbst wenig diskutiert worden. Es wird gleichsam implizit unterstellt, dass die Erhebung und Veröffentlichung der eingeworbenen Drittmittel der Hochschulen sowie der außeruniversitären Forschungseinrichtungen und die Anordnung der Ergebnisse in einer Ranking-Tabelle eine Forschungsleistungsmessung der Institutionen darstellen. Entsprechend wird das Förderranking auch als Forschungsqualitätsranking interpretiert.

Da die DFG als einer der wichtigsten Förderer der Wissenschaft über Informationen der Förderungen unmittelbar selbst verfügt, mag es auch nicht verwunderlich sein, dass sie diese Daten auswertet und veröffentlicht. Man kann aber unterstellen, dass die DFG auch ein Interesse daran hat, Forschungsförderungen als einen Qualitätsnachweis von Forschungen zu interpretieren, weil sie dadurch ihre eigene Legitimität zu stärken vermag. In welchem Maße Drittmittel aber ein geeigneter Indikator zur Messung von Forschungsleistungen sind, werden wir später genauer diskutieren.

Das DFG-Förderranking, das als Forschungsqualitätsranking gedeutet wird, ist für die Reputationszuweisung an die Universitäten von immenser Bedeutung. Gerade weil das „Standing“ der DFG in der Wissenschaft und in der Wissenschaftspolitik fast unumstritten ist, haben ihre Stellungnahmen den Status von sakrosankten Richtersprüchen, an denen sich dann andere Institutionen in ihrer Bewertung von Einrichtungen orientieren. Und in der öffentlichen Rezeption des Förderrankings stößt vor allem eine Tabelle auf die größte Aufmerksamkeit: die Rankingtabelle der Hochschulen. Dies wird deutlich, wenn man sich die Reaktionen der Universitäten am Tag der Veröffentlichung des Forschungsrankings anschaut. Im Folgenden finden sich exemplarisch die Pressemitteilungen der Universitäten Aachen, FU Berlin und TU Dresden nach der Veröffentlichung des letzten Förderrankings:

DFG-Förderatlas weist RWTH Aachen als besonders forschungsstarke Universität aus

24.05.2012

Der heute von der Deutschen Forschungsgemeinschaft (DFG) vorgestellte Förderatlas liefert in umfassenden Analysen Kennzahlen zu der Forschungsaktivität der deutschen Hochschullandschaft. Ein Blick in den Förderatlas zeigt: Mit der Verausgabung von 278 Mio. Euro aus Forschungsmitteln der DFG steht die RWTH Aachen erneut auf dem ersten Platz unter den deutschen Universitäten.

www.rwth-aachen.de/go/id/ciwww

DFG-Förderatlas 2012: Freie Universität unter den ersten Drei. Spitzenreiter in den Geistes- und Sozialwissenschaften

24.05.2012

Im sechsten Förderranking der Deutschen Forschungsgemeinschaft (DFG) hat die Freie Universität Berlin ihre Spitzenposition unter den forschungstärksten Hochschulen bundesweit ausgebaut. Wie die DFG gemeinsam mit der Hochschulrektorenkonferenz (HRK) und dem Stifterverband für die Deutsche Wissenschaft in Berlin mitteilte, verbesserte sich die Freie Universität von Platz fünf im Berichtszeitraum 2005 bis 2007 auf Platz drei im Zeitraum 2008 bis 2010. Ausgewertet wurden die Daten von 186 Hochschulen. Der Rangliste zufolge bewilligte die DFG Wissenschaftlerinnen und Wissenschaftlern der Freien Universität zwischen 2008 und 2010 insgesamt 250,8 Mio. Euro an Fördermitteln. Das sind knapp 30 Prozent mehr als im Vergleichszeitraum 2005 bis 2007. Wenn man die Bewilligungssummen aus der Exzellenzinitiative des Bundes und der Länder nicht berücksichtigt, erreicht die Freie Universität mit eingeworbenen Forschungsmitteln in Höhe von 168,3 Mio. Euro bundesweit Platz eins.

www.fu-berlin.de/presse/informationen/fup/2012/fup_12_134/index.html

Spitzenplätze für die TU Dresden im aktuellen DFG-Förderatlas 2012

25.05.2012

Die Erfolgsgeschichte geht weiter: Im soeben veröffentlichten Förderatlas der Deutschen Forschungsgemeinschaft (DFG) ist die TU Dresden in die Spitzengruppe der deutschen Hochschulen aufgestiegen und erreicht nun Platz 13 – fünf Plätze besser als zuvor. (...)

Im Ranking der DFG-Bewilligungen findet sich die TUD im Berichtszeitraum 2008 – 2010 mit 144 Mio. Euro auf Platz 13. Das führt den Aufwärtstrend der vergangenen 20 Jahre fort: Von Platz 35 im Berichtszeitraum 1991 – 95 über Platz 24 (1996 – 2001), Platz 20 (2002 – 04) und Platz 18 (2005 – 07) bis hin zum 13.

http://tu-dresden.de/aktuelles/newsarchiv/2012/5/dfg_foerderatlas/newsarticle_view

Die Beispiele verdeutlichen, für wie bedeutend die Bewertungen der DFG von den Universitäten interpretiert werden und dass der Rangplatz in der Tabelle der bewilligten Fördermittel als Indikator für die Forschungsstärke einer Universität angesehen wird.

2.2 Die Exzellenzinitiative

Die Exzellenzinitiative des Bundes und der Länder zur Förderung von Wissenschaft und Forschung an deutschen Hochschulen ist in die drei Förderlinien „Graduiertenschulen“, „Exzellenzcluster“ und „Zukunftskonzepte“ gegliedert. Mit der letzten Förderlinie werden ganze Universitäten unterstützt. Die Förderung eines Zukunftskonzeptes setzt die gleichzeitige Förderung von mindestens einem Exzellenzcluster und mindestens einer Graduiertenschule voraus. Für die Reputationszuweisung innerhalb und außerhalb des Wissenschaftssystems ist die dritte Förderlinie deshalb von besonderer Bedeutung. Universitäten, die hier erfolgreich sind, erhalten das Gütesiegel „Exzellenzuniversität“ und gehören damit zu den deutschen Eliteuniversitäten. In der letzten Runde der Exzellenzinitiative, deren Ergebnisse im Juni 2012 publiziert wurden, waren elf Universitäten zu Exzellenzuniversitäten gekrönt worden.

Nach welchen Verfahren wurden die Eliteuniversitäten ausgewählt und welche Kriterien wurden angelegt (vgl. Möller u. a. 2012)? Zuerst zum Verfahren: Im März 2010 haben der Wissenschaftsrat und die DFG die Kriterien für die dritte Runde

der Exzellenzinitiative publiziert. Im ersten Schritt konnten die Universitäten bis zum September 2010 Antragskizzen einreichen. 22 Anträge wurden für die dritte Förderlinie eingereicht. Das Ergebnis der Vorbegutachtung wurde im März 2011 veröffentlicht. Von den 22 Universitäten wurden sieben aufgefordert, einen Vollantrag einzureichen. Diejenigen Universitäten, die in der ersten und zweiten Runde der Exzellenzinitiative erfolgreich gewesen waren, wurden automatisch aufgefordert, Fortsetzungsanträge zu stellen. Über die insgesamt 16 Anträge, die in der Endausscheidung waren, wurde dann in einem mehrstufigen Verfahren geurteilt. **(a)** Die Anträge wurden in einem ersten Schritt von Gutachtergruppen bewertet. Zu dieser Begutachtung gehörte eine Vorstellung und Diskussion des jeweiligen Antrags vor Ort an der jeweiligen Universität. Die Gutachtergruppen haben auf dieser Grundlage eine Empfehlung formuliert. **(b)** Alle Gutachterempfehlungen wurden im nächsten Schritt im Juni 2012 durch eine gemeinsame Kommission aus Vertretern der Deutschen Forschungsgemeinschaft und des Wissenschaftsrates der Bundesregierung bewertet. Die gemeinsame Kommission gab wiederum Empfehlungen an den Bewilligungsausschuss weiter. **(c)** Der Bewilligungsausschuss, in dem neben den Mitgliedern der Kommission auch die für Wissenschaft zuständigen Minister des Bundes und der Länder vertreten sind, hat dann entschieden, dass elf Hochschulen den Status einer Exzellenzuniversität erhalten und bis Ende 2017 gefördert werden sollen (RWTH Aachen, FU Berlin, HU Berlin, Bremen, TU Dresden, Köln, Heidelberg, Konstanz, LMU München, TU München, Tübingen).

Die Kriterien, nach denen die Auswahl in dem mehrstufigen Verfahren faktisch stattfand, sind sehr schwer zu rekonstruieren, weil sie weitgehend in den Arkanbereich der Entscheidungsgremien und der Wissenschaftspolitik fallen und entsprechend von außen nicht einsehbar sind. Am besten nachvollziehbar sind die vom Wissenschaftsrat und der DFG definierten Richtlinien, nach denen der Antrag formuliert sein sollte. Diese Richtlinien geben das Format, die Gliederung, die Gesamtlänge, die Länge der Unterpunkte sowie inhaltliche Gliederungspunkte vor.¹³ So sollten die Antragsteller zuerst den Status Quo der Universität beschreiben, dann das Zukunftskonzept erläutern und dieses schließlich im Rahmen der Langfristplanung der Universität verorten. Die Vorgaben bezüglich der Inhalte des Antrags sind aber sehr allgemein und recht diffus formuliert. Wesentlich konkreter sind die Hinweise, die im Anhang des Antrags aufgelistet werden mussten.

13 www.wissenschaftsrat.de/download/archiv/notes_zwei.pdf

Zu den nachgefragten Informationen gehören u. a.

- Summe der Drittmittel (absolut und pro Professor)
- Platzierung im DFG-Förderranking (2006, 2009)
- SFBs und andere Verbundprojekte
- EU-Mittel
- Bundesmittel, Landesmittel, Mittel aus der Industrie, Stiftungsmittel
- Anzahl der Studierenden, Absolventen, Promotionen, Habilitationen, Berufungen
- Preise

Wie man schnell sieht, wird auch in der Auflistung der Indikatoren, die jenseits der Prosa der Antragsteller die schriftliche Grundlage für die Voten der Gutachter gebildet haben, den Drittmitteln eine außerordentliche Bedeutung zugeschrieben. Und ein besonderer Akzent wird auf die von der DFG vergebenen Fördermittel gelegt.

Wir wissen nicht, in welchem Maße die Drittmittel dann bei der Entscheidungsfindung bedeutsam waren. Leider war dies nicht Bestandteil der Gutachterbefragung von Torger Möller u. a. (2012). Von den elf Universitäten, die am Ende ausgewählt wurden, gehören immerhin sieben zu den ersten elf Hochschulen in der Rangliste der Universitäten nach dem Förderranking der DFG aus dem Jahr 2012 (DFG 2012: 76; Rangliste ohne Mittel der Exzellenzinitiative). Die vier Universitäten, die nicht zu den ersten elf gehören, sind Tübingen auf Platz 15, Köln auf Platz 16, Bremen auf Platz 17 und Konstanz auf Platz 39 von insgesamt 186 Universitäten.

Man kann vermuten, dass die Drittmittelbilanz der Universitäten eine wichtige Rolle auch bei der Auswahl der Exzellenzuniversitäten gespielt hat. Die Antragsprosa, die die Universitäten je für sich verfasst haben, lässt Raum für vielerlei Selbstinszenierungen und Versprechungen, die sich nicht gut kontrollieren lassen. Die schriftlichen Anträge der verschiedenen Bewerber lassen sich zudem sehr schwer miteinander vergleichen. Maßzahlen, die sich in Euro genau ausdrücken lassen, sprechen hingegen eine vermeintlich klare Sprache und erlauben einen unmittelbaren Vergleich zwischen den Universitäten sowie die Bildung von eindeutigen Ranglisten. Entsprechend kann man vermuten, dass in einem Verfahren, in dem es um die relationale Bewertung von Universitäten geht, vergleichende Maßzahlen von wesentlicher Bedeutung sind. Und bei diesem Vergleich spielen

die eingeworbenen Drittmittel wiederum eine ganz entscheidende Rolle. Ob der Indikator der ihm aufgebürdeten Validitätslast auch standhalten kann, werde ich gleich diskutieren.

2.3 Leistungsorientierte Mittelvergabe durch die Bundesländer

Die meisten Hochschulen in Deutschland sind öffentlich durch das jeweilige Bundesland finanziert. Seit den 1990er Jahren wurden von sehr vielen Bundesländern Verfahren der leistungsorientierten Mittelvergabe (kurz: LOM) entwickelt und implementiert (Gläser u. a. 2008; Wissenschaftsrat 2011). Die Länder verfolgen mit der Einführung dieses Instruments vor allem das Ziel, den Wettbewerb zwischen den Universitäten zu befördern und dadurch die Universitäten anzureizen, bessere Leistungen zu generieren, für die sie dann finanziell honoriert werden. Dazu bedarf es der Definition von Kriterien und Indikatoren, mit deren Hilfe die Leistungen der Hochschulen gemessen werden.

Auskunft über die Kriterien, die zur Beurteilung von Leistungen von den Bundesländern angewandt werden, gibt eine Studie von Dominic Orr und Tanja Barthelmes (2010; vgl. auch König 2011: 122f.; Wissenschaftsrat 2011). Die Untersuchung von Orr und Barthelmes fußt auf einer Befragung der Wissenschaftsministerien der sechzehn Länder, die im März 2010 stattgefunden hat. Von diesen haben zwölf geantwortet. Drei Länder vergeben nach eigenem Bekunden keine Mittel nach Leistungskriterien. In die Auswertung der Autoren gehen also neun Länder ein (Baden-Württemberg, Bayern, Berlin, Brandenburg, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen und Rheinland-Pfalz).¹⁴

Die Leistungsbemessung beinhaltet Kriterien, die die Performanz der Universitäten in den beiden Kernbereichen, nämlich Lehre und Forschung, operationalisieren, in manchen Ländern ergänzt um „fachfremde“ Kriterien wie z. B. das Niveau der Frauenförderung. Wir beschränken uns hier auf die Kriterien, die im Bereich der Forschung Leistung messen sollen. In allen Bundesländern werden neben der Anzahl an Promotionen und Habilitationen vor allem und in erster Linie die eingeworbenen Drittmittel zur Messung von Forschungsleistungen

¹⁴ Der Anteil des Budgets, der leistungsorientiert an die Hochschulen vergeben wird, variiert dabei zwischen den Bundesländern beträchtlich. „Spitzenreiter ist Berlin mit einem geplanten Anteil von ca. 30% an den Universitäten (ab 2012); relativ hohe Anteilswerte erreichen auch Nordrhein-Westfalen (rd. 11%), Baden-Württemberg (rd. 10%), Hessen (rd. 10%) und Brandenburg (rd. 8%)“. Einen Überblick über den Anteil an Mitteln, die in anderen Nationen leistungsorientiert vergeben werden, findet man in der Analyse von Diana Hicks (2012: 257).

zugrunde gelegt, Gelder also, die das wissenschaftliche Personal – in der Regel die Professoren – bei den Stiftungen, den Ministerien oder der Industrie einwirbt, um mit diesen finanziellen Mitteln Forschung zu betreiben. Die Länder unterscheiden sich allein in der Art und Weise, wie die Höhe der Drittmittel, die in die Kalkulation eingehen, berechnet werden: als absolute Größe, gewichtet um die Anzahl der Professoren, gewichtet nach Herkunft der Drittmittel, in Referenz zu der jeweiligen Fächergruppe usw. (vgl. hierzu die Tabelle 2 in Orr & Barthelmes 2010). Darüber hinaus honorieren manche Länder zusätzlich die Beteiligung an DFG-Forschungszentren, an Exzellenzclustern oder die Einwerbung von Alexander-von-Humboldt-Stipendien. Ergebnismessungen von Forschungen, z. B. Publikationen, Publikationen in referierten Zeitschriften, Zitationsindizes oder Erfindungen und Patente spielen so gut wie keine Rolle. Eine Ausnahme bildet die Leistungsbemessung in der Medizin in Bayern, bei der Impact-Faktoren von Zeitschriften berücksichtigt werden (Orr & Barthelmes 2010).

Wir sehen also, dass auch bei der Finanzierung der Universitäten nach Leistungskriterien in erster Linie die eingeworbenen Drittmittel als Indikator zur Messung der Forschungsleistung berücksichtigt werden.

Betrachten wir die Situation in Berlin, dem Bundesland, in dem ich selbst tätig bin, etwas genauer (vgl. hierzu Menninghaus 2006). Die Finanzierung der Universitäten setzt sich aus folgenden Posten zusammen: 34,3% der Finanzmittel sind eine Sockelfinanzierung, die leistungsunabhängig ausgezahlt wird. Die restlichen 65,2% der Gelder werden nach definierten Leistungskriterien vergeben. 31,1% werden in die Forschung und die Nachwuchsentwicklung investiert. Die Höhe der Finanzmittel richtet sich nach den erbrachten Leistungen. 29,5% werden für die Lehre vorbehalten. Auch diese Gelder werden nach Leistungskriterien ausgeschüttet. 3,8% werden für Erfolge in der Gleichstellungspolitik und in der sogenannten „Diversity“-Politik eingesetzt (vor allem für Studienanfänger mit Migrationshintergrund). Und schließlich werden 1,3% für Erfolge in der Weiterbildung vergeben.

Da wir uns in unserem Zusammenhang allein für die Messung von Forschungsleistungen interessieren, schauen wir uns die Indikatoren für diesen Bereich genauer an. Nach folgenden Indikatoren werden die Leistungsmittel für Forschung und Nachwuchsentwicklung vergeben:

- 60% für die Summe der eingeworbenen Drittmittel
- 10% für die Anzahl an Publikationen
- 20% für die Anzahl an Promotionen
- 10% für Internationalität, gemessen an der Anzahl an AvH-Stipendiaten

Man sieht, auch hier spielen die Drittmittel in der Bemessung wissenschaftlicher Leistungen eine ganz außergewöhnliche Rolle. Und dies ist unmittelbar für die Universitäten, die Fachbereiche und die Institute handlungsrelevant, weil ihre finanzielle Ausstattung in einem nicht geringen Maße davon abhängt, wie viele Drittmittel sie einwerben. Publikationen, die in anderen Ländern eine wichtige Rolle zur Beurteilung von Forschungsleistungen spielen, sind in Berlin weniger bedeutsam als die Gleichstellungsmaßnahmen. Dass eine solchermaßen von der Politik definierte Anreizstruktur eine große Versuchung darstellt, der die Universitätsleitungen auch in der Rekrutierungspolitik von Professorinnen und Professoren erliegen, mag nicht verwundern. Belastbare Daten zu dieser Vermutung gibt es nicht, Geschichten von Einzelfällen aber hinreichend viele.

Die von einer Universität eingeworbenen Leistungsmittel werden innerhalb der Universitäten nach unten durchgereicht, also je nach Leistung der Fakultäten auf diese und von den Fachbereichen meist auch auf die Institute verteilt. Damit einzelne Professoren sich nicht hinter dem Kollektiv der Universität oder des Fachbereichs wegducken und beim Herstellen eines Kollektivguts in die Rolle des Trittbrettfahrers schlüpfen können, ist die Universitätsleitung dazu übergegangen, in den Berufungszusagen, die mit den neu berufenen Professoren geschlossen werden, das Erfüllen der in den Hochschulverträgen definierten Leistungen festzuhalten. Die Fortsetzung der Zahlung von individuellen Zulagen ist somit an die Erfüllung der festgelegten Leistungen gekoppelt. Nichterfüllung kann also zu Gehaltseinbußen führen. Das verstärkt die Gefolgschaft erheblich.

2.4 Berufungspraxis

Die Fixierung auf Drittmittel als dominantem Indikator zur Leistungsmessung von Forschung spiegelt sich auch in der Berufungspraxis von Hochschullehrern wider. Ohne Anspruch auf Repräsentativität haben wir in einer kleinen Untersuchung analysiert, in welchem Ausmaß der Nachweis von erfolgreicher Drittmittelinwerbung einerseits und von Publikationsleistungen andererseits Bestandteil der Ausschreibung von Stellenanzeigen im meinem eigenen Fach,

der Soziologie, ist. Ausgewertet wurden die Stellenanzeigen für Hochschullehrer (W1, W2, W3) im Fach Soziologie über einen Zeitraum von 18 Monaten (Juni 2010 bis März 2012). Grundlage bildeten die vom Deutschen Hochschullehrerverband wöchentlich an die Mitglieder verschickten Stellenanzeigen. Ausgewählt wurden diejenigen Anzeigen, die im Titel die Widmung Soziologie, Sozialwissenschaften oder empirische Sozialforschung trugen. Wir haben drei Merkmale der Stellenanzeigen ausgewertet. Wird von den Bewerbern erwartet, dass sie **(a)** erfolgreich im Einwerben von Drittmittel waren, **(b)** dass sie in referierten Zeitschriften publiziert haben und/oder **(c)** dass sie unabhängig von der Publikation in Fachzeitschriften sehr gute Publikationen nachweisen können? Insgesamt wurden in dem analysierten Zeitraum 98 Stellenanzeigen veröffentlicht und kodiert.

Wie sehen die Ergebnisse aus? In 46,5% der Fälle gehört es zu den expliziten Anforderungen an die Bewerber, dass sie erfolgreich in der Einwerbung von Drittmitteln waren. Weniger bedeutsam ist offensichtlich die Publikationsleistung. In 27,6% der Anzeigen wird von den Bewerbern verlangt, dass sie in referierten Zeitschriften publiziert haben und in 28,6% der Fälle wird die Erwartung formuliert, dass die Kandidaten sehr gute bzw. exzellente Veröffentlichungen vorweisen können ohne dass spezifiziert wurde, ob es sich dabei um Fachzeitschriften handeln soll oder nicht. Addierte man die Anzahl der Anzeigen, in denen entweder nach Publikationen in Fachzeitschriften oder nach sehr guten Publikationen im Allgemeinen gefragt wurde, kommt man auf einen Wert von 36,7%. Der Einwerbung von Drittmitteln wird also auch hier mehr Bedeutung zugemessen als den Publikationsleistungen der Bewerber.

Die Ausführungen in diesem Abschnitt sollten deutlich gemacht haben, dass in Deutschland den eingeworbenen Drittmitteln im Vergleich zu den Leistungsmessungen in anderen Ländern und im Vergleich zu international vergleichenden Rankings eine besonders zentrale Stellung eingeräumt wird, Publikationen und Zitationen hingegen kaum eine Rolle spielen.

3. WAS TAUGEN DRITTMITTEL ALS INDIKATOR ZUR MESSUNG VON FORSCHUNGSLEISTUNGEN?

In welchem Maße der Indikator Drittmittel eine valide Messung von Forschungsleistungen darstellt und ob der in Deutschland eingeschlagene Weg, den Drittmitteln eine hegemoniale Stellung einzuräumen sinnvoll ist, wird im Folgendem diskutiert.

3.1 Theoretische Kritik: Drittmittel messen den Input und nicht den Output von Forschung

Zur Bestimmung der Leistung einer wissenschaftlichen Einrichtung und des Wissenschaftssystems insgesamt kann man sich an der volkswirtschaftlichen Definition von Produktivität orientieren. Produktivität bezeichnet die Leistungsfähigkeit einer Volkswirtschaft und berechnet sich aus dem Verhältnis von produziertem „Outcome“ z. B. in Form von Gütern oder wissenschaftlichen Erkenntnissen und den dafür benötigten Produktionsfaktoren. Zu den Produktionsfaktoren gehört u. a. das eingesetzte Kapital. Drittmittel in der Forschung sind nun nichts anderes als ein Produktionsfaktor, der Forschung erst ermöglicht, da mit den eingeworbenen Mitteln z. B. Labors eingerichtet und Personal eingestellt werden, die sich dann an die Arbeit machen und Forschungsergebnisse produzieren, die in aller Regel veröffentlicht werden. Drittmittel dienen der Ermöglichung von Forschung, sie sind selbst noch kein Forschungsergebnis. Fördermittel sind in erster Linie eine Input-Größe; die Qualität von Forschung hängt aber ganz entscheidend davon ab, „was hinten raus kommt“. Es hat noch keiner wegen eines hohen Drittmittelaufkommens einen Nobelpreis bekommen, sondern nur wegen guter Erkenntnisse, die sich in Publikationen und Patenten manifestieren.

Man stelle sich einen Automobilkonzern vor, der eine überdurchschnittliche Menge an Kapital oder Arbeitszeit benötigt, um seine Autos zu produzieren. Man käme nicht auf die Idee, die Leistung des Unternehmens am Verbrauch des Kapitals und der Arbeitszeit zur Herstellung von Fahrzeugen zu messen. Ganz im Gegenteil: Das eingesetzte Kapital steht im Nenner des Quotienten der Berechnung der Produktivität. Dies bedeutet entsprechend, dass je weniger Kapital ein Konzern benötigt, um eine bestimmte Menge an Fahrzeugen herzustellen, desto produktiver und leistungsstärker ist er.

Nicht anders verhält es sich mit der Forschung. Entscheidend für die Qualitätseinschätzung ist das Ergebnis von Forschung. Dieses materialisiert sich in Publikationen, in den technischen Fächern zudem in Erfindungen und Patenten. Theoretisch spricht also erst einmal nichts dafür, Drittmittel als Indikator zur Messung von Forschungsleistungen zu benutzen. Sie könnten nur dann einen sinnvollen Indikator zur Messung von Forschungsleistungen darstellen, wenn sie als gute und leicht zugängliche Ersatzmessung für die eigentliche Messgröße dienen.

Von den Befürwortern einer Orientierung an Drittmitteln als Indikator zur Messung von Forschungsleistungen wird nun häufig folgendes Argument ins Feld geführt. Viele der Drittmittelprojekte, vor allem aber diejenigen, die von der DFG und wissenschaftlichen Stiftungen gefördert werden, sind durch ein Bewilligungsverfahren gelaufen. Bestandteil dieses Verfahrens ist in der Regel die Begutachtung des Antrags durch fachkompetente Kolleginnen und Kollegen. Insofern ist die Forschungsqualität zumindest der Projekte, deren Anträge bei Förderinstitutionen eingereicht wurden, die über ein elaboriertes und fachliches Begutachtungsverfahren verfügen, welches dem der Begutachtung von Zeitschriftenartikeln ähnelt, weitgehend gewährleistet (Hornbostel 1997: 212; Neidhardt 2012). Gegen dieses Argument gibt es zwei Gegenargumente. **(a)** Viele der an deutschen Universitäten durchgeführten Drittmittelprojekte haben nicht das beschriebene „strenge“ Begutachtungsverfahren durchlaufen. Der Anteil der Fördersummen von Projekten, die durch die DFG finanziert werden, lag z. B. 2009 bei 35%; 23% der Drittmittelleinnahmen der Hochschulen im Jahr 2009 kamen aus der Industrie und Wirtschaft, 21% vom Bund und die anderen Gelder von Stiftungen und der Europäischen Union. **(b)** Wenn am Ende aus den bewilligten Mitteln keine oder qualitativ schlechte Forschungsergebnisse entstehen, nützt eine gute Bewertung bei der Vergabe der Mittel zur Bestimmung der Qualität der Forschung nur wenig. Insofern sind die Drittmittel als Indikator zur Messung der Forschungsqualität nur dann plausibel, wenn sie mit einem guten Forschungsergebnis einhergehen. Dies leitet zum nächsten Punkt über.

3.2 Empirische Kritik: Der statistische Zusammenhang zwischen Drittmitteln und Ergebnismessungen von Forschungen

Angenommen, jeder, der mit dem H1N1-Virus infiziert ist, der zur sogenannten Schweinegrippe führt, weist auf dem linken Ohrläppchen einen kleinen grünen Punkt auf. In diesem Fall könnte man sich aufwendige Virusuntersuchungen sparen, um eine Schweinegrippeninfektion zu diagnostizieren. Der kleine grüne Punkt wäre ein guter Indikator für den H1N1-Virus. Da nur sehr wenige Leute, die nicht infiziert sind, einen kleinen grünen Punkt am linken Ohrläppchen haben, wäre der korrelative Zusammenhang zwischen dem Ohrläppchenmerkmal und einer Schweinegrippeninfektion fast perfekt. Auch wenn der grüne Punkt selbst nichts über den H1N1-Virus und seine Beschaffenheit aussagt, könnte man ihn als Ersatzmessung zur Diagnose dieser Infektion benutzen. Mit einem ähnlichen Argument ließe sich theoretisch die Messung von Forschungsleistungen durch die Erhebung von Drittmitteln legitimieren. Ähnlich wie der grüne Punkt bei der Diagnose einer Krankheit hat der Drittmittelindikator viele pragmatische Vorteile. Die Daten über die eingeworbenen Drittmittel sind sehr gut verfügbar, weil sie auf allen Ebenen des Forschungssystems bereits erhoben werden; zudem handelt es sich um eine metrische Größe, die es erlaubt, klare Rankingtabellen zu erstellen. Die entscheidende Frage ist aber, ob Drittmittel wirklich die Qualität von Forschungsleistungen anzeigen, in welchem Maße sie also mit Indikatoren, die das Resultat von Forschungen messen, korrelieren.

Schauen wir uns im Folgenden genauer an, in welchem Maße die eingeworbenen Drittmittel mit den erläuterten „Outcome“-Messungen von Forschungen korrelieren. Versucht man diesbezüglich den Forschungsstand zu bilanzieren, wird man zuerst feststellen müssen, dass es relativ wenige Studien gibt, die diesen Zusammenhang systematisch untersucht haben. Angesichts der Tatsache, dass die Drittmittel eine so bedeutsame Rolle in der Forschungsbewertung und in der Forschungsförderung spielen, ist dies ein erstaunlicher Befund. Mittlerweile gibt es eine Vielzahl von Personen und Institutionen, die in der Wissenschaftsforschung engagiert sind. In Deutschland ist mit dem Institut für Forschungsinformation und Qualitätssicherung (IFQ) hierzu eine eigene Einrichtung geschaffen worden. Dies hat nicht dazu geführt, dass der Zusammenhang zwischen Drittmittelaufkommen und anderen Performanzindikatoren systematisch untersucht wurde. Hinzu kommt, dass die vorliegenden Studien nicht immer den methodischen Kriterien guter Analyse entsprechen. Ich werde im Folgenden einige der Ergebnisse verschiedener Studien berichten.

- a) Stefan Hornbostel (2001) hat den Zusammenhang zwischen Drittmitteln und Publikationen für die Fächer Mathematik, Physik, Informatik und Recht analysiert. Die Daten zur Messung der Drittmittel und der Publikationen beruhen auf der Selbstauskunft der befragten Professoren. Für die verschiedenen Fächer ergeben sich folgende Korrelationskoeffizienten. Physik: .50; Mathematik: .40; Informatik: .33; Recht: .31.¹⁵ Die DFG hat in ihrem letzten Förderranking (2012: 177) für das Fach Chemie den Zusammenhang zwischen Drittmitteln und Publikationen berechnet. Die Analyse bezieht sich allerdings nur auf die 40 Universitäten mit dem höchsten Drittmittelaufkommen. Der Koeffizient erreicht den hohen Wert von .71.
- b) René Krempkow, Uta Landrock und Patricia Schulz (2012: 252) haben jüngst den Zusammenhang zwischen Drittmitteln und Publikationen im Bereich Medizin analysiert. Der Korrelationskoeffizient zwischen den Drittmitteln pro Professur (2003-2005) und den Publikationen pro Professur (2006-2008) der von ihnen untersuchten Fachbereiche beträgt .35.
- c) Deutlich schwächer sind die Zusammenhänge im Bereich der Geistes- und Sozialwissenschaften. Auf der Basis einer Sekundäranalyse der Daten des CHE (Centrum für Hochschulentwicklung) zeigt Hornbostel (2005) für die Erziehungswissenschaften, dass die Korrelation zwischen Drittmitteln und Publikationen .045 beträgt und nicht signifikant ist. Für die Soziologie berechnet Hornbostel einen Wert von .271. Berücksichtigt man bei der Drittmittelberechnung allein die DFG-Mittel, steigt der Wert für die Erziehungswissenschaften auf .316. Es gibt also einen positiven Zusammenhang zwischen den Drittmitteln und der Anzahl der Publikationen, dieser scheint aber für die Erziehungswissenschaften sehr schwach zu sein.
- d) Dieses Ergebnis wird für die Soziologie durch eine Studie des Wissenschaftsrats (2008) bestätigt. Der Wissenschaftsrat hat in einer Pilotstudie unter der Leitung von Friedhelm Neidhardt mit einem sehr aufwendigen Verfahren das Fach Soziologie flächendeckend einer Evaluation unterzogen und in diesem Kontext auch verschiedene Indikatoren zur Messung von Forschungsleistungen ausprobiert. Die Korrelation zwischen der Menge der publizierten Artikel in so genannten „Peer-Review“-Zeitschriften und

15 Für die Physik hat Hornbostel noch eine zweite Berechnung durchgeführt, die sich nicht auf Publikationen des Professors allein bezieht, sondern auf die jeweilige Arbeitsgruppe und die Publikationen der Arbeitsgruppe in Zeitschriften, die im „Science Citation Index“ gelistet sind. Der korrelative Zusammenhang fällt dann etwas höher aus.

den eingeworbenen Drittmitteln beträgt demnach .218. Dabei gehen in die Analyse nur die DFG-Drittmittel ein. Wenn man die Mittel aus der Wirtschaft und Industrie hinzunähme, die unter Umständen nach anderen, auf jeden Fall zusätzlichen Kriterien als die Mittelvergabe bei der DFG vergeben werden, wäre der Wert wahrscheinlich wesentlich niedriger.

- e) Das Centrum für Hochschulforschung hat bis 2009 ein Forschungsranking erstellt. In das Ranking und in die Profilbeschreibung der verschiedenen Fächer gehen u. a. folgende Indikatoren ein: Drittmittel, Anzahl der Promotionen, Publikationen, Zitationen und Erfindungsmeldungen. Zudem wurde die Forschungsreputation der Universitäten in den untersuchten Fachgebieten auf der Grundlage einer Befragung erhoben. Das CHE-Forschungsranking ist ab 2011 in der Studie „Vielfältige Exzellenz“ aufgegangen. Hier werden jetzt zusätzlich zu den erwähnten Indikatoren noch andere Dimensionen und Indikatoren wie der Anwendungsbezug, die Internationalität und die Studierendenorientierung der Fächer analysiert.
- f) Richard Münch (2006) hat den Zusammenhang zwischen Drittmitteln und verschiedenen Publikationswerten mit Hilfe des Forschungsrankings des CHE aus dem Jahr 2005 tabellarisch dargestellt. Die folgenden beiden Tabellen aktualisieren die Ergebnisse von Münch und beziehen sich auf das Forschungsranking des CHE aus dem Jahr 2009 (Tabelle 3) und auf die Nachfolgestudie „Vielfältige Exzellenz“ aus dem Jahr 2011 (Tabelle 4). Die Informationen wurden aus den Berichten der beiden Studien des CHE zusammengestellt. Dort finden sich auch die Details zur Methodik (Berghoff u. a. 2009, 2011).

Tabelle 3: Zusammenhang zwischen Drittmitteln je Wissenschaftler und „Outcome“-Indikatoren von Forschungsleistungen sortiert nach Korrelationsstärke (angegeben sind Korrelationskoeffizienten): Bezug: Berghoff u. a. (2009).¹⁶

Fächer	Publikationen pro Wissenschaftler	International sichtbare Publikationen pro Wissenschaftler ^e	Erfindungen pro Wissenschaftler	Zitationen pro Publikation ^f
Pharmazie	0,56		0,37 ^a	29,01%
Zahnmedizin	0,56			19,49%
Biologie	0,49**		0,11	18,25%
Medizin	0,41		0,26	6,44%
Psychologie	0,38*			6,70%
ET & IT ^d	0,35*		0,34	5,11%
Anglistik	0,26			14,99%
Geschichte	0,24			
Chemie	0,21		0,23 ^a	0,18
Erziehungswiss.	0,2			
Soziologie	0,175 ^a			
BWL	-0,03	0,29* ^a		
VWL	-0,02	0,06 ^a		
Maschinenbau			0,7**	
Physik			0,18 ^a	0,16

** Die Korrelation ist auf dem Niveau von 0,01 zweiseitig signifikant.

* Die Korrelation ist auf dem Niveau von 0,05 zweiseitig signifikant.

a) Hier wurde nicht Publikationen pro Wissenschaftler, sondern pro Professor erhoben.

b) Hier wurde nicht internationale Publikationen pro Wissenschaftler, sondern pro Professor erhoben.

c) Hier wurden nicht Erfindungen pro Wissenschaftler, sondern pro zehn Wissenschaftler erhoben.

d) Elektrotechnik und Informationstechnik.

e) Unter international sichtbaren Publikationen versteht das CHE Publikationen des Typus „Journal Article“, „Review Article“ und „Letter“, die in den Datenbanken SCI/SSCI/A&HCI des ISI Web of Science gelistet sind.

f) Grundlage bildet eine bibliometrische Analyse der Datensätze SCI, SSCI und A&HCI des Web of Science.

Tabelle 4: Zusammenhang zwischen Drittmitteln je Wissenschaftler und „Outcome“-Indikatoren von Forschungsleistungen sortiert nach Korrelationsstärke (angegeben sind Korrelationskoeffizienten): Bezug: Berghoff u. a. (2011).¹⁷

Fächer	Publikationen pro Wissenschaftler	International sichtbare Publikationen pro Wissenschaftler ^e	Erfindungen pro Wissenschaftler	Zitationen pro Publikation ^f
Psychologie	0,42			0,28
ET & IT ^d	0,32		0,22 ^c	
Anglistik / Amerikanistik	0,26			
Erziehungswiss.	0,26			
BWL	0,01	0,24*		
VWL	-0,02	0,36* ^b		
Maschinenbau	0,26		0,37 ^c	

Dass die meisten Korrelationszusammenhänge statistisch nicht signifikant sind, geht vor allem auf die geringe Fallzahl zurück. Inhaltlich sind die folgenden Befunde besonders interessant.

- a) Zuerst einmal kann man sich darüber freuen, dass es in fast keinem Fach einen negativen Zusammenhang zwischen den eingeworbenen Drittmitteln einerseits und den verschiedenen Messungen der Ergebnisse von Forschung andererseits gibt.
- b) Je nach „Outcome“-Messung fällt der korrelative Zusammenhang unterschiedlich aus. Er ist am stärksten, wenn man die Anzahl der Publikationen berücksichtigt. Aber selbst hier werden keine – von Ausnahmen abgesehen – sehr hohen Werte erreicht. Hinzu kommt, dass die schiere Anzahl an Publikationen selbst kein guter Indikator für die Qualität von Publikationen ist.

17 www.che.de/downloads/CHE_AP149_Vielfaeltige_Exzellenz_2011_2_Auflage.pdf

- c) Deutlich schwächer sind die korrelativen Zusammenhänge für Outcome-Messungen, in denen die eigentliche Qualität der wissenschaftlichen Leistung in der ein oder anderen Weise geprüft wurde (Patente, Publikationen in begutachteten Zeitschriften, Zitationen). Hier variieren die Korrelationskoeffizienten zwischen .37 und .06.
- d) Auch wenn die Tabelle nahe legt, dass Drittmittel in keinem Fach eine optimale Operationalisierung der Ergebnisse von Forschungen darstellen, gilt dies für die verschiedenen Fächer in einem unterschiedlichen Ausmaß. Gruppieren man die Fächer entlang einer Skala, die von den Natur- und Technikwissenschaften bis hin zu den Geisteswissenschaften reicht, dann eignen sich Drittmittel in der zuerst genannten Wissenschaftsgruppe noch am ehesten zur Messung von Forschungsleistungen. Die Ursache für diesen Fächerunterschied liegt u. a. im Typus der Forschung. Natur- und vor allem technikwissenschaftliche Forschung bedarf neben dem Personal meist einer Apparateausstattung, ohne die eine Forschung gar nicht möglich ist (Hornbostel 2001). Diese Ausstattung muss zumindest partiell über Drittmittel finanziert werden (Jansen u. a. 2007: 132). In den Geisteswissenschaften sind hingegen häufig eine gute Bibliothek und der Zugang zu Archiven, die als öffentliche Güter bereits zur Verfügung stehen und nicht gesondert beantragt werden müssen, eine hinreichende Voraussetzung, um mit der Forschung beginnen zu können. Sozialwissenschaftliche Theoretiker oder Geisteswissenschaftler brauchen z. B. zum Schreiben von größeren Monographien vor allem Forschungszeit. Niklas Luhmann schreibt diesbezüglich in der Einleitung zu seinem Opus Magnum einer Gesellschaftstheorie: „Bei meiner Aufnahme in die 1969 gegründete Fakultät für Soziologie der Universität Bielefeld fand ich mich konfrontiert mit der Aufforderung Forschungsprojekte zu benennen, an denen ich arbeite. Mein Projekt lautete: Theorie der Gesellschaft; Laufzeit: 30 Jahre; Kosten: keine“ (Luhmann 1997: 11).¹⁸

Insgesamt scheinen die wenigen Studien, die den Anteil von eingeworbenen Drittmitteln mit anderen, theoretisch plausibleren Messungen von Forschungsergebnissen in Beziehung setzen, den Schluss nahe zu legen, dass Drittmittel kein sehr guter Indikator zur Messung von Forschungsleistungen sind. Umso

18 Auch innerhalb der Geistes- und Sozialwissenschaften gibt es Forschungsfragen, die drittmittelintensiv sind. Dazu gehören zum Beispiel Editionsprojekte. Umgekehrt gilt für die Naturwissenschaften, dass je nach Fach der Drittmittelbedarf unterschiedlich ausfällt. Mathematiker stellen z. B. deutlich seltener Projektanträge als Physiker (Hornbostel 2001: 529).

erstaunlicher ist es, dass die wissenschaftlichen und wissenschaftspolitischen Institutionen in Deutschland fast unbekümmert an den Drittmitteln als Indikator zur Messung von Forschungsleistungen festhalten.

Allerdings muss man in Rechnung stellen, dass fast alle Studien, die bis dato vorliegen, Korrelationen zwischen Drittmitteln einerseits und Publikationsmaßen andererseits auf der Ebene von Instituten oder Fachbereichen berechnen. Wenn man von der Ebene der Institutionen aber auf das Verhalten von Individuen zurück schließt, kann es sein, dass man einem ökologischen Fehlschluss aufsitzt. Denn ein korrelativer Zusammenhang auf der Aggregatsebene muss nicht notwendigerweise einen Zusammenhang auf der Individualebene abbilden. Um solche möglichen Fehlschlüsse aufzudecken, aber auch um den Zusammenhang zwischen Drittmitteln einerseits und Publikationen und Zitationen andererseits systematisch zu untersuchen, ist dringend mehr Forschung notwendig.

3.3 Kritik an der Art und Weise der Berechnung und Nutzung des Indikators Drittmittel

Von einer grundsätzlichen Kritik an der Verwendung von Drittmitteln als Indikator zur Messung von Forschungsleistungen lassen sich Kritikpunkte unterscheiden, die sich auf die spezifische Verwendung des Indikators durch wissenschaftliche Institutionen beziehen.

- a) Wie gerade erläutert, sind die verschiedenen Fächer in unterschiedlichem Maße von der Einwerbung von Drittmitteln abhängig, um überhaupt Forschungen bewerkstelligen zu können. Und auch die Summe der benötigten Drittmittel variiert je nach Fach gewaltig. Eine beantragte Ausstattung mit technischen Apparaten ist eine ganz andere Finanzgröße als die Beantragung einer halben wissenschaftlichen Mitarbeiterstelle. „Ein Projekt von einer Million Euro [...] muss nicht um einen Faktor zehn besser sein, als ein Projekt, bei dem es nur um den Zuschlag von 100.000 Euro geht. Beantragte und bewilligte Geldmittel sind ein Ausdruck des Geldbedarfs von Forschung, und dieser Bedarf variiert vor allem mit den Merkmalen des Projekts, seiner Laufzeit und danach, in welcher Disziplin und in welchem ihrer Forschungsfelder dieses Projekt laufen soll“ (Neidhardt 2012). Die *systematischen* Bedarfsunterschiede der Fächer haben unmittelbare Folgen für die Ranking-Tabellen, wenn sich diese allein auf die absoluten

Summen der eingeworbenen Drittmittel beziehen (Münch 2007). Technische Universitäten und Fachbereiche, die medizinisch oder eher technisch-naturwissenschaftlich orientiert sind, haben einen deutlich höheren Drittmittelbedarf als Universitäten und Fachbereiche, die einen geistes- und sozialwissenschaftlichen Schwerpunkt haben. Kontrolliert man nun alle anderen Einflussfaktoren, schneiden entsprechend die naturwissenschaftlich-technisch ausgerichteten Universitäten und Fachbereiche in der Regel in den drittmittelbasierten Rankings deutlich besser ab als die geistes- und sozialwissenschaftlichen Universitäten und Fachbereiche.

Eine recht einfache und leicht durchzuführende Korrektur dieser Verzerrung besteht darin, nicht (oder nicht nur) die Summe der eingeworbenen Drittmittel auszuweisen, sondern die Anzahl der bewilligten Projekte und dies unabhängig von der Bewilligungssumme (Neidhardt 2012) darzustellen. Sonderforschungsbereiche könnte man dabei z. B. durch die Anzahl der bewilligten Einzelprojekte berücksichtigen. In dem wirkungsmächtigen DFG-Ranking, aber auch in anderen Rangfolgetabellen, fehlen solche Maßzahlen bislang völlig, obwohl sie wahrscheinlich deutlich aussagekräftiger wären.

- b) Angenommen drei Bauarbeiter, die einen Lastwagen mit Sand entladen und diesen mit Schubkarren abtransportieren, benötigen für diese Tätigkeit sechs Stunden Zeit. An einer anderen Baustelle benötigt ein Bauarbeiter für dieselbe Tätigkeit zwölf Stunden. Vergleicht man die für die Tätigkeit benötigte Zeit miteinander, dann ist der LKW im ersten Fall doppelt so schnell entladen worden wie im zweiten Fall. Berücksichtigt man hingegen die benötigte Zeit in Relation zum Personaleinsatz, dann ist der zweite LKW deutlich schneller entladen worden. Angesichts der Trivialität des an diesem Beispiel illustrierten Prinzips ist man erstaunt, dass Ranking-Tabellen sehr häufig allein auf absoluten Zahlen beruhen und eine Relationierung um den Personaleinsatz nicht – oder nur an zweiter Stelle – ausgewiesen wird (zum Folgenden vgl. Münch 2007). Eine Universität oder eine außeruniversitäre Forschungseinrichtung, die über eine große Anzahl an wissenschaftlichem Personal verfügt, ist in der Einwerbung von Drittmitteln, aber auch im Hinblick auf andere Indikatoren (Anzahl der eingeworbenen Alexander-von-Humboldt-Stipendien, Anzahl der DFG-Gutachter usw.) im Vergleich zu einer Einrichtung mit relativ wenig wissenschaftlichem Personal – *ceteris paribus* – erfolgreicher. Richard Münch hat in seiner Studie (2007) die Orientierung an absoluten statt an relationalen Maßzahlen vehement kritisiert. Er geht

davon aus, dass ein solches Vorgehen entscheidend zur Oligopolbildung im deutschen Wissenschaftssystem beiträgt, da die großen Universitäten systematisch bevorteilt werden (zur Kritik an Münch vgl. Auspurg u. a. 2008; Mayer 2009; Hornborstel 2009).

Schauen wir uns unter dem Gesichtspunkt der relationalen Messung von Forschungsleistungen den DFG-Förderatlas aus dem Jahr 2012 genauer an. Im Hinblick auf die Drittmittel weist die DFG zwar an manchen Stellen auch die personalrelativierten DFG-Bewilligungen aus. Dies gilt vor allem dann, wenn die einzelnen Fächerguppen separat betrachtet werden. Addiert man die Anzahl an Tabellen des Berichts, in denen absolute Größen und personalrelativierte Größen ausgewiesen werden, dann zeigt sich aber, dass ganz dominant Tabellen mit absoluten Größen dargestellt werden.

Der DFG-Förderatlas enthält aber nicht nur Informationen über die eingeworbenen Drittmittel, sondern weist auch andere Leistungsindikatoren aus. Auch für diese gilt, dass eine personalrelativierte Größe wesentlich aussagekräftiger ist als absolute Maßzahlen. So ist z. B. die Anzahl der DFG-Gutachter, die von einer Universität oder einer außeruniversitären Einrichtung kommen, abhängig von der Anzahl der Professoren, die an einer Einrichtung beschäftigt sind. Große Universitäten haben eine deutlich höhere Chance Gutachter zu stellen als kleine Universitäten. Folgende Indikatoren werden von der DFG neben den Drittmitteln in den Förderatlas (2012) aufgenommen (in Klammern ist jeweils die Seitenzahl angegeben, auf der sich die jeweiligen Tabellen im Bericht der DFG befinden).

- Mitglieder der DFG-Fachkollegien pro Einrichtung (S. 67)
- Mitglieder der DFG-Gutachterinnen und -Gutachter pro Einrichtung (S. 67, 117, 127, 141, 157)
- Aufenthalte von Geförderten der Alexander-von-Humboldt-Stiftung pro Einrichtung (S. 67, 118, 128, 140, 156)
- Förderungen durch einen „European Research Grant“ pro Einrichtung (S. 67)

In *allen* Fällen werden allein absolute Zahlen und nicht um das wissenschaftliche Personal gewichtete Maßzahlen tabellarisch dargestellt. Die Tatsache, dass die DFG fast nur absolute Größen und nicht um den Personaleinsatz gewichtete Zahlen publiziert, zeichnet ein verzerrtes Bild.

Dieses Bild wird aber gerade von den großen Institutionen dankend aufgenommen und für die Selbstinszenierung und Eigenwerbung benutzt.

- c) Tabellen und Berichte, die Forschungsleistungen von wissenschaftlichen Einrichtungen durch die Summe der eingeworbenen Drittmittel operationalisieren, unterstellen implizit, dass das Verhältnis zwischen Drittmitteln und wissenschaftlichen Leistungen z. B. in Form von Publikationen eine lineare Funktion ist. Je mehr Drittmittel eine Institution eingeworben hat, desto besser ist vermeintlich ihre Forschungsleistung. Die Wissenschaftsforschung zeigt nun allerdings, dass der statistische Zusammenhang zwischen der Menge der eingeworbenen Drittmittel und der Menge der Publikationen nicht nur schwach ausgeprägt ist, sondern kurvilinear zu verlaufen scheint (vgl. Jansen u. a. 2007). Ab einer gewissen Anzahl von Drittmitteln leidet die wissenschaftliche Produktivität. Dorothea Jansen und Mitautoren (2007) haben für die Bereiche Astrophysik, Nanotechnologie und Mikroökonomie den Einfluss der Drittmittel auf die Anzahl an Publikationen untersucht und kommen zu folgendem Ergebnis: „Drittmittel wirken demzufolge zunächst positiv auf den Output, ab einem bestimmten kritischen Niveau aber negativ“ (Jansen u. a. 2007: 137). Zu viele Drittmittel, so die Vermutung, führen dazu, dass die Verantwortlichen zu Drittmittelmanagern degenerieren und nicht mehr zum Publizieren kommen.

3.4 Ein Sonderweg in der Leistungsmessung ist nicht sinnvoll

Dass Drittmittel nach jetzigem Kenntnisstand allein kein sehr guter Indikator zur Messung von Forschungsleistungen sind, sollten die Ausführungen in den letzten Abschnitten gezeigt haben. Damit verschärft sich aber das Problem, dass Deutschland bei der Bestimmung der Leistungsindikatoren einen eigenen, einen Sonderweg geht. Bewertungskriterien von wissenschaftlichen Leistungen sind nicht nur harmlose Orientierungsgrößen, sie wirken auf das Wissenschaftsfeld zurück. Die Einführung von Evaluationskriterien wirkt verhaltenssteuernd, da sich die Akteure im Feld in ihren Handlungen an der Erfüllung der spezifischen Kriterien orientieren, ganz unabhängig davon, ob sie die Kriterien nun als sinnvoll erachten oder nicht (Gläser u. a. 2008; Butler 2010; Orr & Barthelmes 2010). Jochen Gläser u. a. (2008) zeigen für Australien, dass mit der Definition von Drittmitteln als Merkmal der Leistungsmessung auch thematische Folgen verbunden sind, insofern diejenigen Forschungsgebiete in den Vordergrund rücken, die

drittmittelintensiv sind. Henk Moed u. a. (1998) stellen für flämische Universitäten fest, dass sich mit einer Zunahme der externen Finanzierung der Forschung im Zeitverlauf der Abstand zwischen forschungsstarken und forschungsschwachen Universitäten verstärkt.

Wenn gilt, dass die Definition von Gütekriterien von Forschung das Verhalten der Akteure in der Weise beeinflusst, dass sie ihre Zeit darauf verwenden, den Gütekriterien nachzukommen, und wenn weiterhin gilt, dass der Zusammenhang zwischen Drittmitteln einerseits und Publikationen und Zitationen andererseits kein sehr starker Zusammenhang ist, dann werden im deutschen Wissenschaftssystem mit seiner Drittmittelfixiertheit nicht die richtigen Anreize gesetzt. In internationalen Vergleichen wird ein Mehr an Drittmitteln der deutschen Forschung nur wenig nutzen, da sich hier die Bewertung von Forschungsleistungen in erster Linie an unterschiedlichen Messungen von Publikationen und deren Zitationen orientiert, wie wir oben erläutert haben. Will das deutsche Wissenschaftssystem im internationalen Leistungsvergleich besser abschneiden, müssen Publikationen und Zitationen zu den zentralen Indikatoren der Leistungsbemessung werden.

Darüber hinaus gibt es sehr konkrete Maßnahmen, die man ergreifen kann, um die gemessene Leistungsbilanz in internationalen Rankings zu verbessern. Auf zwei dieser Maßnahmen sei hier kurz eingegangen.

- a) **Symbolische Integration der außeruniversitären Forschungseinrichtungen:** Die meisten Bewertungen der Forschungsleistungen beziehen sich allein auf die Leistungen, die in Universitäten erbracht werden. Das deutsche Wissenschaftssystem ist aber bekanntlich durch eine duale Struktur von universitärer und außeruniversitärer Forschung (Max-Planck-Gesellschaft, Fraunhofer Gesellschaft, Helmholtz-Gemeinschaft, Leibniz-Gemeinschaft) gekennzeichnet (Mayer 2012). Forschungsleistungen, die an den außeruniversitären Instituten erbracht werden, gehen in die Leistungsbemessung der Universitäten und damit in die Gesamtbilanz der Bemessung der Forschungsleistung Deutschlands nicht ein. Das Problem ist lange bekannt und vom Wissenschaftsrat in verschiedenen Kontexten thematisiert worden. Nun ist es sicherlich illusorisch, eine Fusion der außeruniversitären Einrichtungen mit den Universitäten zu erwarten. Und es bleibt auch dahin gestellt, ob eine solche Fusion überhaupt sinnvoll ist. Allerdings könnte man über eine symbolische Doppelmitgliedschaft der Mitglieder der

außeruniversitären Forschungseinrichtungen nachdenken. Dies hätte zur Folge, dass die Wissenschaftler der außeruniversitären Einrichtungen bei ihren Publikationen eine doppelte institutionelle Affiliation angeben würden, die der außeruniversitären Einrichtung und die der affilierten Universität, was wiederum dazu führen würde, dass die Publikationen der außeruniversitären Einrichtungen in die Rankings der deutschen Universitäten eingehen und den Wissenschaftsstandort Deutschland nominell deutlich stärken würden.

- b) Förderung englischsprachiger Publikationen: Derselbe wissenschaftliche Text mit den gleichen inhaltlichen Forschungsergebnissen hat je nach Sprache, in der er publiziert wird, eine unterschiedliche Chance, von anderen Forschern im Feld wahrgenommen und dann auch zitiert zu werden. Je höher die Anzahl der Sprecher einer Sprache (Muttersprachler und Sprecher, die diese Sprache als Fremdsprache sprechen), desto höher ist die Wahrscheinlichkeit, dass ein wissenschaftlicher Text gelesen wird. Und da die Lektüre in aller Regel die Voraussetzung für eine Zitation ist, steigt mit dem Verbreitungsgrad einer Sprache auch die Zitationswahrscheinlichkeit. Und da die Zitationsraten von Artikeln wiederum eine zentrale Messung der Qualität von Forschung darstellen, die in fast allen Rankings benutzt werden, entscheidet die Publikationssprache auch über den Platz eines Landes oder einer Universität innerhalb eines Rankings. Aus hier nicht weiter zu diskutierenden Gründen hat die Anzahl der Publikationen in Englisch im Zeitverlauf kontinuierlich zugenommen, so dass Englisch zur dominanten Wissenschaftssprache geworden ist (Gerhards 2010).

Den Zusammenhang zwischen der Publikationssprache und der Zitationswahrscheinlichkeit hat Werner Marx (2001) sehr schön an einem Beispiel illustriert, indem er die Zitationsraten von Artikeln im Fach Chemie, die zugleich auf deutsch und auf englisch publiziert wurden, miteinander verglichen hat. Die englische Version wurde ca. viermal häufiger zitiert als die deutschsprachige Variante, obwohl der Inhalt der Artikel identisch war (vgl. auch Van Leeuwen u. a. 2001). Da Englisch die hegemoniale Sprache ist, sind Zitationen von englischsprachigen Publikationen wahrscheinlicher als von Veröffentlichungen in anderen Sprachen.

Während in den Natur- und Technikwissenschaften die Hürde des Publizierens auf Englisch nicht mehr besonders hoch zu sein scheint, gilt dies nicht für die

Geistes- und Sozialwissenschaften. Hier wird die überwiegende Anzahl der Publikationen weiterhin auf Deutsch veröffentlicht. Damit wird aber der Anschluss an die internationale Kommunikation und die Zitationswahrscheinlichkeit deutlich erschwert. Die fördernden Forschungseinrichtungen in Deutschland täten gut daran, einen Teil ihres Etats für Übersetzungen zur Verfügung zu stellen. Bei Projektbewilligungen könnte man eine Teilsumme für Übersetzungen reservieren und die Projektnehmer verpflichten, zumindest einen Teil ihrer Ergebnisse auf Englisch zu publizieren. Geht man von einer konservativen Berechnung aus und unterstellt, dass die Zitation eines englischen Textes im Durchschnitt doppelt so hoch ist wie die eines deutschen Textes, dann wäre die Förderung von Übersetzungen eine sehr effektive Mittelverwendung. Man kann sich kaum eine Maßnahme vorstellen, wie man die Zitationsraten mit gleichem Mitteleinsatz ähnlich gut verbessern könnte.

4. ZUSAMMENFASSUNG UND AUSBLICK

Die systematische Messung von Forschungsleistungen hat seit den 1990er Jahren weltweit zunehmend an Bedeutung gewonnen. Man kann davon ausgehen, dass auch zukünftig dem indikatorenbasierten Vergleich sowohl im internationalen als auch im nationalen und inneruniversitären Kontext eine zentrale Rolle zukommt. Entsprechend bedeutsam ist es, ob und in welchem Maße die benutzten Indikatoren die jeweilige Forschungsqualität valide messen.

Die Summe der eingeworbenen Drittmittel bildet in der Bundesrepublik die zentrale Messgröße zur Bestimmung der Qualität von Forschung. Die Zentralakteure des deutschen Wissenschaftssystems, vor allem die Deutsche Forschungsgemeinschaft, aber auch der Wissenschaftsrat, haben einen entscheidenden Anteil daran, dass Drittmittel als Forschungsleistungsindikatoren so hoch gehalten, wenn nicht fetischisiert werden. So ist das Forschungsranking der DFG, das in regelmäßigen Abständen publiziert wird und das enorm wirkungsmächtig ist, fast ausschließlich drittmittelbasiert.

Der Wissenschaftsrat (2011) hat sich jüngst in seinen Empfehlungen zur Bewertung und Steuerung von Forschungsleistungen mit den verschiedenen, indikatorenbasierten Verfahren der Leistungsmessung auseinandergesetzt. Zwar warnt der Wissenschaftsrat in seiner sehr allgemein gehaltenen Stellungnahme davor, die Nutzung von Indikatoren zur Leistungsmessung zu verabsolutieren

und empfiehlt, bei der Forschungsbewertung ein den einzelnen Fächern angemessenes „informed peer review“-Verfahren einzusetzen, in dem eine Gutachtergruppe auf der Basis verschiedener Indikatoren die Leistungen beurteilen soll. Er hat es aber versäumt, die besondere Bedeutung, die Drittmitteln in Deutschland zugesprochen wird, zu betonen und vor allem den Drittmittel-Indikator in Abgrenzung zu alternativen Messungen von Forschungsleistungen einer kritischen Betrachtung zu unterziehen.

Das Hauptziel der Ausführungen war es, zu diskutieren, ob Drittmittel ein geeigneter Indikator zur Messung der Qualität von Forschung sind und ob der in Euro darstellbaren Messgröße eine solch dominante Stellung eingeräumt werden sollte, wie dies in Deutschland der Fall ist. Es hat den Anschein, dass Drittmittel aus theoretischen wie empirischen Gründen nicht sonderlich gut geeignet sind, die Qualität von Forschung zu messen. Zumindest sollte ihnen nicht die hegemoniale Stellung in der Beurteilung von Forschungsqualität eingeräumt werden. Theoretisch gilt, dass Drittmittel per definitionem ein Produktionsfaktor der Ermöglichung von Forschung sind, aber nicht die Produktivität und das Ergebnis von Forschungen messen. Das Kernziel von wissenschaftlicher Forschung besteht in der Produktion von Erkenntnisfortschritt. Dieser materialisiert sich in Publikationen und Erfindungen und deren positiver Bewertung durch „peers“. Eine Qualitätsevaluation von Forschung muss hier ansetzen. Eine Ersatzmessung, die sich allein oder ganz dominant auf Drittmittel stützt, ist nur dann legitim, wenn der korrelative Zusammenhang zwischen Drittmitteln und den Resultaten von Forschungen sehr hoch ist. Zwar ist der methodisch gesicherte Kenntnisstand zu dieser Frage eher dürftig; die bis dato vorliegenden Ergebnisse zeigen aber, dass der statistische Zusammenhang nicht besonders ausgeprägt ist.

Ich hatte in der Einleitung bereits die hier formulierte Kritik an der deutschen Drittmittelfixiertheit in mehreren Punkten eingeschränkt und möchte am Ende darauf nochmals zurückkommen.

- a) Es gibt verschiedene Dimensionen, entlang derer man Forschung bewerten kann. Der Wissenschaftsrat (2008) unterscheidet in seinem Forschungsrating u. a. zwischen „Forschungsqualität“ einerseits und „Transfer von Forschung in andere gesellschaftliche Bereiche“ andererseits. Die hier vorgetragene Kritik bezieht sich ausschließlich auf die Nutzung von Drittmitteln als Indikator zur Messung der Forschungsqualität und nicht auf die Nutzung von Drittmitteln zur Messung der Dimension „Transfer“. Ein Teil der Forschung

ist bekanntlich anwendungsorientierte Forschung, die auf die Lösung von konkreten, praktischen Problemen ausgerichtet ist. Zur Operationalisierung dieser Dimension mögen Drittmittel ein guter Indikator sein.

- b) Die formulierte Kritik bezieht sich auf die Nutzung des Drittmittelaufkommens als dominantem Indikator zur Messung der Qualität von Forschung. Mit der Kritik ist nicht gemeint, dass Drittmittel für bestimmte Fächer und in Kombination mit publikationsbasierten Indikatoren nicht auch zur Messung der Forschungsqualität genutzt werden können. Die berichteten Korrelationen zwischen Drittmitteln einerseits und Publikationen und Zitationen andererseits fallen aber für die Geistes- und Sozialwissenschaften am niedrigsten und für die technischen und naturwissenschaftlichen Fächer am höchsten aus.
- c) Wenn man Drittmittel zur Messung der Forschungsqualität heranzieht, scheinen mir mehrere, zum Teil einfach umzusetzende Korrekturen notwendig zu sein. Zum einen sollten die eingeworbenen Drittmittel nicht als absolute Größe, sondern im Verhältnis zum wissenschaftlichen Personal berücksichtigt werden. Zum zweiten ist es ratsam, nicht die Drittmittelsummen, sondern die Anzahl der eingeworbenen Projekte auszuweisen, um dem sehr unterschiedlichen Drittmittelbedarf der Fächer gerecht zu werden. Drittens muss man bei der Nutzung von Drittmitteln als Indikator zur Messung der Forschungsqualität berücksichtigen, ob die Bewilligung der Projekte durch ein nach wissenschaftlichen Kriterien strukturiertes Peer-Review-Verfahren erfolgt ist oder nicht.

Wir wissen, dass die Definition der Gütekriterien von Forschung das Verhalten der Akteure beeinflusst. Wenn gilt, dass der Zusammenhang zwischen Drittmitteln einerseits und Publikationen und Zitationen andererseits kein sehr starker ist, dann werden in der Bundesrepublik mit einer übermäßigen Hochachtung, ja Fetischisierung der Drittmittel die falschen Anreize gesetzt. In internationalen Vergleichen wird ein Mehr an Drittmitteln der deutschen Forschung nur wenig nutzen, da sich hier die Bewertung von Forschungsleistungen in erster Linie an unterschiedlichen Messungen von Publikationen und deren Zitationen orientiert, wie wir oben erläutert haben. Der deutsche Sonderweg könnte sich langfristig als wissenschaftspolitische Sackgasse erweisen.

Nun mag man einwenden, dass ich die alternativen Qualitätsmessungen, zu denen es eine breite und kritische Diskussion gibt, zu wenig betrachtet habe. Und in der Tat ist eine Leistungsmessung, die sich an Publikationen und Zitationen orientiert, alles andere als unproblematisch (vgl. dazu zum Beispiel Committee on Quantitative Assessment of Research 2008; Glänzel 2008; Marx 2011; Moed 2005; Persson u. a. 2004; Weingart & Winterhager 1994; Weingart u. a. 1991; Winterhager & Weingart 1997). Es muss einem anderen Papier vorbehalten bleiben, die Kritikpunkte, die man gegenüber publikationsbasierten Verfahren der Qualitätsbewertung anbringen kann, zu diskutieren.

Es gibt aber drei zentrale Argumente, warum ein an Publikationen orientiertes Verfahren der Leistungsmessung besser ist als eine Drittmittelorientierung – und man zumindest eine kombinierte Betrachtung anraten muss. Erstens liegen Publikationen und Zitationen als empirisch messbare Phänomene deutlich näher an dem theoretischen Konstrukt „Forschungsleistung“. Zweitens lassen sich nach meiner Ansicht die Probleme, die mit bibliometrischen Verfahren der Leistungsmessung verbunden sind, zwar nicht lösen, aber leichter verbessern als eine Orientierung an Drittmitteln. Drittens steht eine Orientierung an Publikationen im Einklang mit vielen international vergleichend angelegten Messungen von Forschungsleistungen. Eine bessere Institutionalisierung bibliometrischer Verfahren in Deutschland würde die Anreize so setzen, dass sie im Einklang mit internationalen Standards stehen, was langfristig zu einem besseren Abschneiden der deutschen Forschung in internationalen Vergleichen führen kann. Winston Churchill hatte in einer Rede im Unterhaus (am 11. November 1947) die Vorteile einer Demokratie mit folgenden Worten beschrieben: „Demokratie ist die schlechteste aller Regierungsformen – abgesehen von all den anderen Formen, die von Zeit zu Zeit ausprobiert worden sind.“ In Abwandlung dieses Zitats würde ich behaupten, dass bibliometrische Verfahren die schlechtesten aller Verfahren zur Forschungsleistungsmessung sind, abgesehen von allen anderen Formen, die im Moment vor allem in Deutschland institutionalisiert sind bzw. ausprobiert werden.

LITERATUR

Auspurg, Katrin, Thomas Hinz & Jürgen Güdler. 2008. Herausbildung einer akademischen Elite? Zum Einfluss der Größe und Reputation von Universitäten auf Forschungsförderung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60, S. 653–685

Berghoff, Sonja, Gero Federkeil, Petra Giebisch, Cort-Denis Hachmeister, Mareike Hennings, Isabel Roessler & Frank Ziegele. 2009. Das Hochschulranking 2009. Vorgehensweise und Indikatoren. Gütersloh. www.che-ranking.de/downloads/CHE_AP119_Methode_Hochschulranking_2009.pdf

Berghoff, Sonja, Petra Giebisch, Cort-Denis Hachmeister, Britta Hoffmann-Kobert, Mareike Hennings & Frank Ziegele. 2011. Vielfältige Exzellenz 2011: Forschung - Anwendungsbezug - Internationalität - Studierendensorientierung im CHE Ranking. Gütersloh. www.che-ranking.de/downloads/CHE_AP149_Vielfaeltige_Exzellenz_2011_2_Auflage.pdf

Botzem, Sebastian. 2012. *The Politics of Accounting Regulation. Organizing Transnational Standard Setting in Financial Reporting*. Cheltenham/Northampton, MA: Verlag Edward Elgar

Butler, Linda. 2010. Impacts of Performance-based Research Funding Systems: A Review of the Concerns and the Evidence. In: OECD (Hg.): *Performance-based Funding for Public Research in Tertiary Education Institutions. Workshop Proceedings*. www.oecd-ilibrary.org/education/performance-based-funding-for-public-research-in-tertiary-education-institutions/impacts-of-performance-based-research-funding-systems-a-review-of-the-concerns-and-the-evidence_9789264094611-7-en

Committee on Quantitative Assessment of Research. Citation Statistics. 2008. A Report from the International Mathematical Union (IMU) in Cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf

Esser, Hartmut. 1999. *Soziologie. Spezielle Grundlagen. Band 1: Situationslogik und Handeln*. Frankfurt am Main: Campus

Gaethgens, Peter. 2012. Die Exzellenzinitiative im Kontext Bund/Länder-finanzierter Forschungsförderprogramme. *Wissenschaftspolitik im Dialog* 1/2012. Berlin-Brandenburgische Akademie der Wissenschaften

Gerhards, Jürgen. 2002. Reputation in der deutschen Soziologie. *Zwei Welten. Soziologie* 31(2), S. 19–33

Gerhards, Jürgen. 2010. Mehrsprachigkeit im vereinten Europa. *Transnationales sprachliches Kapital als Ressource in einer globalisierten Welt*. Wiesbaden: Verlag VS für Sozialwissenschaften

- Glänzel, Wolfgang & András Schubert.** 2004. Analyzing Scientific Networks Through Co-Authorship. In: Henk F. Moed, Wolfgang Glänzel & Ulrich Schmoch (Hg.): Handbook of Quantitative Science and Technology Research. The Use of Publication and Patent Statistics in Studies of S&T Systems. Dordrecht: Kluwer Academic Publishers, S. 257–276
- Glänzel, Wolfgang.** 2008. Seven Myths in Bibliometrics. About Facts and Fiction in Quantitative Science Studies. In: Proceedings of WIS 2008, Berlin, Humboldt-Universität zu Berlin, Institute for Library and Information Science (IBI), S. 24–32
- Gläser, Jochen.** 2008. Evaluationsbasierte Managementsysteme für universitäre Forschung. BMBF-Studie. www.bmbf.de/pubRD/evaluationsbasiertes_managementsystem_universitaere_forschungsleistung.pdf
- Gläser, Jochen, Thomas H. Spurling & Linda Butler.** 2004. Intraorganisational Evaluation: Are there „Least Evaluable Units? Research Evaluation 13(1), S. 19–32
- Gläser, Jochen, Stefan Lang, Grit Laudel & Uwe Schimank.** 2008. Evaluationsbasierte Forschungsfinanzierung und ihre Folgen. In: Renate Mayntz, Friedhelm Neidhardt, Peter Weingart & Ulrich Wengenroth (Hg.): Wissensproduktion und Wissenstransfer: Wissen im Spannungsfeld von Wissenschaft, Politik und Öffentlichkeit. Bielefeld: Verlag transcript, S. 145–170
- Hicks, Diana.** 2012. Performance-based University Research Funding Systems. Research Policy 41, S. 251–261
- Hornbostel, Stefan.** 1997. Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag
- Hornbostel, Stefan.** 2001. Third Party Funding of German Universities. An Indicator of Research Activity? Scientometrics 50, S. 523–537
- Hornbostel, Stefan & Edwin Keiner.** 2002. Evaluation der Erziehungswissenschaft. Zeitschrift für Erziehungswissenschaft 4, S. 634–653
- Jansen, Dorothea, Andreas Wald, Karola Franke, Ulrich Schmoch & Torben Schubert.** 2007. Drittmittel als Performanzindikator der wissenschaftlichen Forschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59(1), S. 125–149
- Kocka, Jürgen.** 1988. German History before Hitler: The Debate about the German "Sonderweg". Journal of Contemporary History 23(1), S. 3–16
- König, Karsten.** 2011. Hochschulsteuerung. In: Peer Pasternack (Hg.): Hochschulen nach der Föderalismusreform. Leipzig: Akademische Verlagsanstalt, S. 106–154
- Krempkow, René, Uta Landrock & Patricia Schulz.** 2012. Steuerung durch LOM? Eine Analyse zur leistungsorientierten Mittelvergabe an medizinischen Fakultäten in Deutschland. In: Uwe Wilkesmann & Christian Schmidt (Hg.): Hochschule als Organisation. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 245–260
- Krück, Andreas.** 2011. Private Ratings, Public Regulations. Credit Rating Agencies and Global Financial Governance. Basingstoke: Palgrave Macmillan

- Kühl, Stefan.** 2013. Drittmittel für Universitäten. Entzauberung des Fetischs. Süddeutsche Zeitung. 4. Januar 2013
- Laudel, Grit.** 2006. The „Quality Myth“: Promoting and Hinderling Conditions for Acquiring Research Funds. *Higher Education* 52, S. 375–403
- Lindenberg, Siegwart.** 1996. Die Relevanz theoriereicher Brückenannahmen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 48, S. 126–140
- Luhmann, Niklas.** 1972. Selbststeuerung der Wissenschaft. In: Niklas Luhmann, *Soziologische Aufklärung*. Opladen: Westdeutscher Verlag., S. 232–252
- Luhmann, Niklas.** 1997. *Die Gesellschaft der Gesellschaft*. Frankfurt a.M.: Suhrkamp Verlag
- Marx, Werner.** 2001. Angewandte Chemie in Light of the Science Citation Index. *Angewandte Chemie – International Edition* 40(1), S. 139–143
- Marx, Werner.** 2011. Bibliometrie in der Forschungsbewertung. *Forschung & Lehre* 18(11), S. 858–860
- Mayer, Karl Ulrich.** 2009. Der Geist hat die Universität verlassen. *Soziologische Revue* 32: 3–8
- Mayer, Karl Ulrich.** 2012. Produktive Pfadabhängigkeiten. Ein Diskussionsbeitrag zum Verhältnis universitärer und außeruniversitärer Forschung im Kontext der Exzellenzinitiative. *Wissenschaftspolitik im Dialog* 3/2012. Berlin-Brandenburgische Akademie der Wissenschaften
- Merton, Robert K.** 1968. The Matthew Effect in Science. *Science* 159(3810), S. 56–63
- Moed, Henk F.** 2005. *Citation Analysis in Research Evaluation*. Dordrecht: Springer Verlag
- Moed, Henk F., Marc Luwel, Josee A. Houben, Eric Spruyt & Herman Van Den Berghe.** 1998. The Effects of Changes in the Funding Structure of the Flemish Universities on their Research Capacity, Productivity and Impact during the 1980's and early 1990's. *Scientometrics* 43, S. 231–255
- Münc, Richard.** 2006. Drittmittel und Publikationen. *Forschung zwischen Normalwissenschaft und Innovation*. *Soziologie* 35(4), S. 440–461
- Münc, Richard.** 2007. *Die akademische Elite. Zur sozialen Konstruktion wissenschaftlicher Exzellenz*. Frankfurt a.M.: Suhrkamp
- Neidhardt, Friedhelm.** 1988. *Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG*. Opladen: Westdeutscher Verlag
- Neidhardt, Friedhelm.** 2008. Das Forschungsrating des Wissenschaftsrats. Einige Erfahrungen und Befunde. *Soziologie* 37(4), S. 421–432
- Neidhardt, Friedhelm.** 2010. *Selbststeuerung der Wissenschaft: Peer Review*. In: Dagmar Simon, Andreas Knie & Stefan Hornbostel (Hg.): *Handbuch Wissenschaftspolitik*. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 280–293
- Neidhardt, Friedhelm.** 2012. *Maßstäbe kritischer Würdigung von Wissenschaft. Zum 15-jährigen Bestehen des Bielefelder Instituts für Konflikt- und Gewaltforschung*. Berlin: Unveröffentlichtes Vortragsmanuskript

- OECD.** 2010. Performance-based Funding for Public Research in Tertiary Education Institutions. Workshop Proceedings. www.oecd-ilibrary.org/education/performance-based-funding-for-public-research-in-tertiary-education-institutions_9789264094611-en
- Orr, Dominic & Tanja Barthelmes.** 2010. Leistungsorientierte Finanzierung öffentlicher Forschung im tertiären Bildungssektor in Deutschland. Hannover: HIS Arbeitspapier
- Persson, Olle, Wolfgang Glänzel & Rickard Danell.** 2004. Inflationary Bibliometric Values. The Role of Scientific Collaboration and the Need for Relative Indicators in Evaluative Studies. *Scientometrics* 60(3), S. 421–432
- Power, Michael.** 1999. *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press
- Stichweh, Rudolf.** 1988. Differenzierung des Wissenschaftssystems. In: Renate Mayntz, Bernd Rosewitz, Uwe Schimank & Rudolf Stichweg (Hg.): *Differenzierung und Verselbständigung. Zur Entwicklung gesellschaftlicher Teilsysteme*. Frankfurt a.M.: Campus Verlag, S. 11–44
- Möller, Torger, Philipp Antony, Sybille Hinze & Stefan Hornbostel.** 2012. Exzellenz begutachtet. Befragung der Gutachter in der Exzellenzinitiative. iFQ-Working Paper No.11. Berlin
- Van Leeuwen, Thed N., Henk F. Moed, Robert J.W. Tussen, Martijn S. Visser & Anthony F.J. Van Raan.** 2001. Language Biases in the Coverage of the Science Citation Index and its Consequences for International Comparisons of National Research Performance. *Scientometrics* 51, S. 335–346
- Weingart, Peter & Matthias Winterhager.** 1984. *Die Vermessung der Forschung. Theorie und Praxis der Wissenschaftsindikatoren*. Frankfurt a. M./New York: Campus Verlag
- Weingart, Peter, Wolfgang Prinz, Maria Kastner, Sabine Maasen & Wolfgang Walter.** 1991. *Die sogenannten Geisteswissenschaften. Außenansichten*. Frankfurt a. M.: Suhrkamp Verlag
- Weingart, Peter.** 2001. *Die Stunde der Wahrheit? Zum Verhältnis der Wissenschaft zu Politik, Wirtschaft und Medien in der Wissensgesellschaft*. Weilerswist: Velbrück Verlag
- Winterhager, Matthias & Peter Weingart.** 1997. *Forschungsstatus Schweiz 1995*. Schweizerischer Wissenschaftsrat (Hg.): Reihe Forschungspolitik, Bern, FOP 45/1997
- Wissenschaftsrat.** 2008. *Pilotstudie Forschungsrating. Empfehlungen und Dokumentation*. Köln: Wissenschaftsrat. www.wissenschaftsrat.de/download/Forschungsrating/Dokumente/FAQ/Pilotstudie_Forschungsrating_2008.pdf
- Wissenschaftsrat.** 2011. *Empfehlungen zur Bewertung und Steuerung von Forschungsleistungen*. Halle: Wissenschaftsrat. www.wissenschaftsrat.de/download/archiv/1656-11.pdf

WEITERE TITEL DER REIHE »WISSENSCHAFTSPOLITIK IM DIALOG«

Heft 1/2012

PETER GAEHTGENS

Die Exzellenzinitiative im Kontext Bund/Länder-finanzierter
Forschungsförderprogramme

Heft 2/2012

HANS MEYER

Die Zukunft des Wissenschaftssystems und die Regeln des Grundgesetzes über
Sach- und Finanzierungskompetenzen

Heft 3/2012

KARL ULRICH MAYER

Produktive Pfadabhängigkeiten.

Ein Diskussionsbeitrag zum Verhältnis universitärer und außeruniversitärer
Forschung im Kontext der Exzellenzinitiative

Heft 4/2012

STEPHAN LEIBFRIED/ULRICH SCHREITERER

Quo vadis, Exzellenzinitiative?

Heft 5/2012

REINHARD HOFFMANN

Das monistische Modell.

Die Mitfinanzierung des Bundes von Universitätseinrichtungen des
Landes im integrativen Forschungsverbund Universität/außeruniversitäre
Forschungseinrichtung

Heft 6/2013

HANS-GERHARD HUSUNG

Zukunftsraum Wissenschaft.

Was kommt nach der Exzellenzinitiative?

Die vier wichtigsten Bund-Länder-Programme für die Forschung (Hochschulpakt, Exzellenzinitiative, Hochschulbau, Pakt für Forschung und Innovation) enden zwischen 2013 und 2020, so dass in den nächsten Jahren wichtige politische Entscheidungen für die zukünftige Gestaltung der deutschen Forschungslandschaft gefällt werden müssen. Die Schriftenreihe *Wissenschaftspolitik im Dialog* bietet ein Forum für Analysen der bisherigen Instrumente der Wissenschafts- und Forschungsförderung wie auch für eine breit gefächerte offene Diskussion über die Zukunft des deutschen Wissenschaftssystems. Die Reihe wird von der interdisziplinären Arbeitsgruppe Exzellenzinitiative der BBAW betreut.