

Zur Zukunft von Forschungsratings

Vier Thesen der Jungen Akademie



Das deutsche Wissenschaftssystem ist in den vergangenen Jahren auf bislang ungekannte Weise von einer „Evaluitis“ befallen worden. Evaluationen bilden ein Kernelement des New Public Management, das an den Universitäten Einzug gehalten hat. Die Budgethöhe der Hochschulen, die leistungsorientierte Mittelzuweisung und die W-Besoldung mit ihren Zielvereinbarungen und Leistungskontrollen haben die klassischen Universitäten in Unternehmen verwandelt, die eigene Interessen an der Steigerung ihrer Produktivität und ihres Kapitals entwickeln.

Die Dramatik dieser Entwicklung ist hinlänglich bekannt.¹ Die Junge Akademie hat etliche damit verbundene Erscheinungen, wie etwa die Orientierung an drittmittelfinanzierter Verbundforschung, bereits nachdrücklich kritisiert und auf die im Vergleich zur ex-

post-Evaluation viel höhere Bedeutung guter Berufungsverfahren aufmerksam gemacht.² Dass die Weichenstellung in Richtung rein betriebswirtschaftlicher Managementmethoden für viele Bereiche der Wissenschaft problematisch ist, bleibt eine Überzeugung der Jungen Akademie. Gerade wenn der Zug einmal abgefahren ist, ergibt sich aber die Frage: Was tun? Die „Pilotstudie Forschungsrating“ des Wissenschaftsrats, deren erste Ergebnisse in diesen Tagen veröffentlicht werden, bietet eine willkommene Gelegenheit, sich mit diesem Problem zu befassen.

In seiner Pilotstudie hat der Wissenschaftsrat einige vielfach kritisierte Schwächen von Evaluationen überwunden. Zwei Einsichten zeichnen das exemplarisch für die beiden Fächer Chemie und Soziologie durchgeführte Forschungsrating aus. Erstens: Das Rating ist *multidimensional* angelegt; es erfasst verschiedene Aspekte der Qualität, des Impacts und der Effizienz von Forschung ebenso wie Nachwuchsförderung und Wissenstransfer. Zweitens: Das Rating ist *disziplinspezifisch* angelegt; die für das „informed peer review“ verantwortlichen Bewertungsgruppen haben je eigene quantitative und qualitative Indikatoren für die genannten Dimensionen gebildet. Schon jetzt scheint sich abzuzeichnen, dass

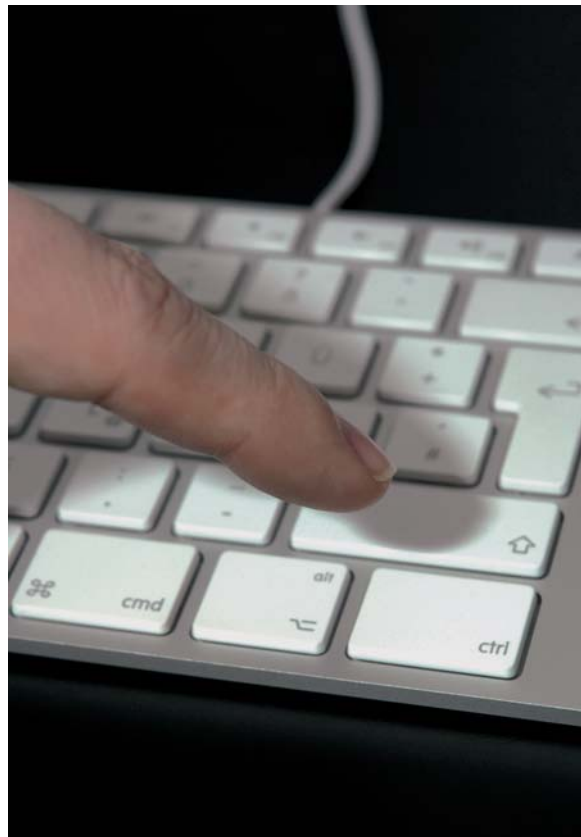
¹ Bruno Frey, „Evaluitis – eine neue Krankheit“, Working Paper No. 293, Institute for Empirical Research in Economics, University of Zürich, 2006; ähnlich der Beitrag von Margit Osterloh und Bruno Frey in der Frankfurter Allgemeinen Zeitung vom 21. Juli 2007; vgl. auch Richard Münch, Die akademische Elite, Frankfurt am Main: Suhrkamp 2007.

² Vgl. das Thesenpapier der AG Manieren!, „Wissenschaft als Betrieb und Norm(al)fabrik. Fünf Anmerkungen zur Wissenschaftsförderung“, Berlin, 2006; sowie Giovanni Galizia/Julia Fischer, „Konkurrenz und Kooperation. Zum Thema Berufungsverfahren in Deutschland“, Berlin, 2005.

dieses Forschungsrating eine Selbstverständigung der betroffenen Fächer über Standards angeregt hat. Nachdem die Ergebnisse der Pilotstudie vorliegen, stellt sich für den Wissenschaftsrat allerdings die Frage, ob Forschungsratings nun auch für andere Fächer konzipiert und in regelmäßigen Abständen durchgeführt werden sollten. Die folgenden Thesen verstehen sich als Beitrag zur Diskussion um diese Frage.

1. Evaluationen sollten Qualität, nicht Quantität beurteilen

Bisherige Evaluationen begnügen sich damit, das zu messen, was leicht messbar ist. Eingeworbene Drittmittel gehören ebenso dazu wie bibliometrische Angaben über die Anzahl von Publikationen, Zitationshäufigkeit etc. Diese Kennzahlen sind nicht nur weit von dem entfernt, was für den wissenschaftlichen Erkenntnisfortschritt eigentlich zählt, nämlich kreative Ideen, originelle Theorien und innovative Methoden. Sie schaffen auch – und dies wiegt noch schwerer – falsche Anreizstrukturen. So werden WissenschaftlerInnen ihre Forschungsprojekte wegen höherer Bewilligungschancen am Mainstream orientieren oder zur Verbesserung ihres Gehalts mehr publizieren, ohne dass dies zu nennenswerten Erkenntnisfortschritten beiträgt.³ Dies betrifft häufig auch NachwuchswissenschaftlerInnen, die sich nahezu permanent beurteilen lassen müssen. Wir wünschen uns daher Evaluationen, die sich an der inhaltlichen Qualität veröffentlichter Forschungsergebnisse orientieren und auch weiterhin das „Bohren dicker Bretter“ lohnend erscheinen lassen. Die in der Pilotstudie von der Bewertungsgruppe Soziologie gewählte Praxis, die jeweils besten Veröffentlichungen begutachten zu lassen, ist zwar aufwendig, liegt aber näher an den innerhalb einer scientific community geltenden Qualitätsstandards als die Kennziffern eingeworbener Drittmittel.



2. Evaluationen müssen disziplinspezifisch sein

Die bislang üblichen Rankings krankten an starren Kriterienkatalogen. Erkennbarer Erkenntnisfortschritt lässt sich jedoch nicht in allen Fächern gleichermaßen an Drittmitteln oder publizierten Artikeln in begutachteten Fachzeitschriften ablesen; in etlichen Geisteswissenschaften bleiben Monographien die wichtigsten Beiträge zum jeweiligen Fachdiskurs. Gerade wenn Hochschulleitungen, die in den seltensten Fällen einen kompetenten Überblick über das gesamte Fächerspektrum ihrer Universität besitzen, vernünftige und nachvollziehbare Entscheidungen über die Vergabe von Mitteln treffen wollen, bedürfen sie fachspezifischer, qualitätsbezogener Informationen statt allgemeiner und wenig aussagekräftiger Kennziffern. Wir möchten den Wissenschaftsrat daher unbedingt darin bestärken, den in der Pilotstudie gegangenen Weg fortzusetzen und disziplinspezifische Indikatoren zu bilden. Für einzelne Fächer, wie beispielsweise die Philosophie, könnte dies auch bedeuten, auf Drittmittel als Indikator vollständig zu verzichten.

³ Empirische Evidenz für diese Effekte lässt sich in Australien finden, wo die Einführung von monetären Anreizen für Publikationen zu einer Flut irrelevanter Veröffentlichungen führte; vgl. Linda Butler, „A list of published papers is no measure of value“, *Nature* 419 (2002): 877



3. Evaluationen sollten vorhandene Ressourcen berücksichtigen

Die Leistung von Forschungseinheiten wird in vielen Rankings an absoluten Kennziffern gemessen. Gerade bei quantitativen Indikatoren bedeutet dies eine deutliche Verzerrung zugunsten überdurchschnittlich gut ausgestatteter Institutionen, insbesondere großer Universitäten und außeruniversitärer Forschungseinrichtungen wie der Max-Planck-, Helmholtz- und Fraunhofer-Institute. Diese Unterschiede werden insbesondere dann noch verstärkt, wenn der zeitliche Einsatz in der grundständigen Lehre und damit der elementaren Ausbildung des wissenschaftlichen Nachwuchses unberücksichtigt bleibt. Die Aufnahme von Bewertungskriterien der Effizienz in der Pilotstudie des Wissenschaftsrats ist daher zu begrüßen. Wir fordern, dass zukünftige Evaluationen die (wie immer gemessenen) Forschungsleistungen stets in ein angemessenes Verhältnis zu den vorhandenen personellen, finanziellen und zeitlichen Möglichkeiten setzen; dies ist im Übrigen auch für die Familienverträglichkeit wissenschaftlicher Karrieren relevant.

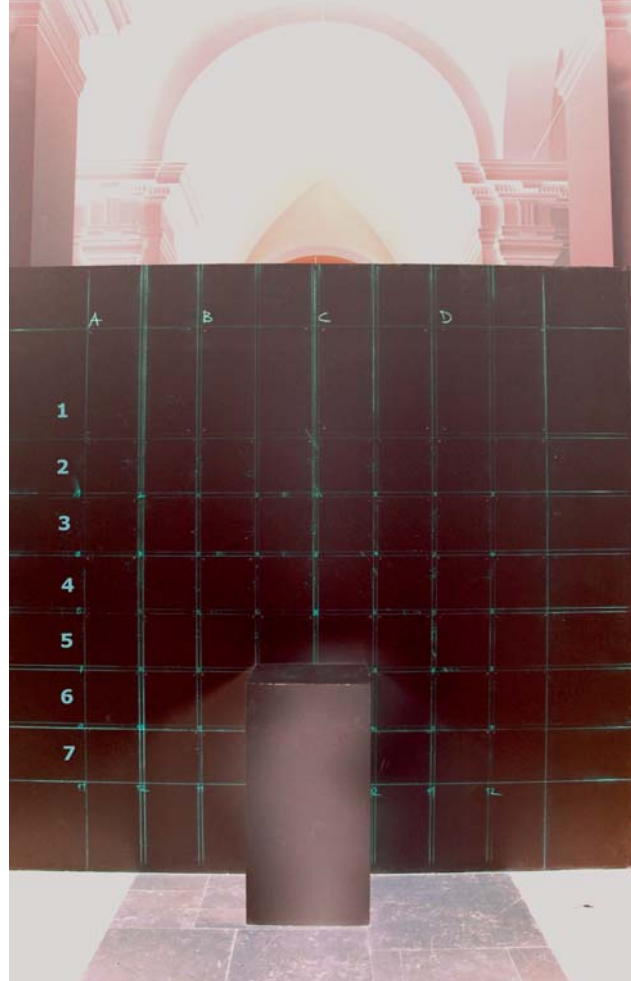
4. Evaluationen sollten schlank sein und selten erfolgen

Der Aufwand für Evaluationen der Forschungseinheiten ganzer Fächer steht in keinem Verhältnis zu dem gewonnenen Informationsgewinn. Es werden enorm viel Zeit und Geld auf die Datenerhebung, die Datenauswertung und die Erstellung von Gutachten und Evaluationsberichten verwendet. Die Evaluierten selbst sowie die Gutachterinnen und Gutachter verlieren dadurch Zeit für die eigene Forschung. Der Wissenschaftsrat sollte die Pilotstudie nutzen, um zu prüfen, welche Indikatoren sich besonders gut für die fachspezifische Evaluation der Qualität von Forschung eignen und welche eher irrelevant sind. Nachfolgende Evaluationen könnten vor diesem Hintergrund schlanker ausfallen. Sie sollten, zumal angesichts der ambivalenten Erfahrungen mit dem Research Assessment Exercise in Großbritannien, auch möglichst sel-

ten stattfinden. Und sie sollten stets Rechenschaft über das Verhältnis von Kosten und (vermutetem) Nutzen abgeben.

Nur wenn man diesen vier Vorschlägen folgt, werden Forschungsratings die Informationsbasis von Universitätsleitungen und Ministerien und damit (hoffentlich) auch deren Entscheidungen über die Verteilung von Mitteln innerhalb und zwischen Universitäten verbessern können. Würden sich Evaluationen streng an innerfachlichen Qualitätskriterien und institutionellen Voraussetzungen von Kreativität orientieren und sich nicht dem Diktat messbarer Kennziffern unterwerfen, könnten sie dazu beitragen, dass die skurril anmutende Phase der Selbstverstümmelung akademischer Freiheit beendet und der Begriff guter Wissenschaft wieder dorthin gegeben wird, wo er hingehört, nämlich in die Republik des Geistes.

- Bettina Beer, Michael Boutros, Julia Fischer, Matthias Klatt, Matthias Koenig, Thomas Koop, Melanie Schnell, Christine Silberhorn



Zur Zukunft von Forschungsratings veranstaltet die Junge Akademie am 21. April 2008 in Berlin eine öffentliche Podiumsdiskussion (siehe auch S. 26).

Ansprechpartner: Matthias Koenig & Thomas Koop