

# Die Vermessung der Wissenschaft

Messung und Beurteilung  
von Qualität in der Forschung

Wien, im Juli 2014

ÖSTERREICH  
WISSENS

## **Die Vermessung der Wissenschaft**

### **Messung und Beurteilung von Qualität in der Forschung**

## Inhalt

Executive Summary .....	3
Einleitung.....	6
1. Mess- und Beurteilungsverfahren in internationaler Perspektive.....	9
2. Mess- und Beurteilungsverfahren in nationaler Perspektive .....	15
2.1 Verfahren an österreichischen Universitäten.....	15
2.2 Verfahren an außeruniversitären Einrichtungen.....	21
2.2.1 Beispiel <i>Fonds zur Förderung der wissenschaftlichen Forschung</i> (FWF).....	21
2.2.2 Beispiel <i>Institute for Science and Technology</i> (IST Austria).....	25
2.2.3 Beispiel <i>Österreichische Akademie der Wissenschaften</i> (ÖAW).....	27
3. Stellungnahme und Empfehlungen .....	31
3.1 Allgemeine Einschätzungen .....	31
3.2 Österreichspezifische Empfehlungen .....	38
3.2.1 Im Allgemeinen .....	39
3.2.2 Im Besonderen.....	41

*„An der Akademie hielt Humboldt Vorlesungen über die Leitfähigkeit menschlicher Nerven. Er stand dabei, als im Nieselregen auf ausgetretenem Rasen vor der Stadt der letzte Abschnitt des Längengrades gemessen wurde, der Paris mit dem Pol verband. Als es vollbracht war, nahmen alle die Hüte ab und schüttelten einander die Hände: Ein Zehnmillionstel der Strecke würde, in Metall gefasst, zur Einheit der künftigen Längenmessungen werden. Man wollte es Meter nennen. Es erfüllte Humboldt stets mit Hochgefühl, wenn etwas gemessen wurde; diesmal war er trunken vor Enthusiasmus. Die Erregung ließ ihn mehrere Nächte nicht schlafen.“<sup>1</sup>*

## **Executive Summary**

Die Wissenschaft ist es gewohnt zu messen. Sie ist es ebenso gewohnt, selbst gemessen zu werden, Rechenschaft über ihr Tun abzulegen. Diese ‚Vermessung‘ kann Wege der Forschung und ihre Erträge sichtbar machen und Anliegen einer wissenschafts- und forschungspolitischen Steuerung dienen; sie kann Basis zukünftiger Schwerpunktsetzungen, Anreiz zur Profilbildung einer forschenden Institution als Bildungsanbieter sein. Sie kann der Selbsteinschätzung von Fächern und Disziplinen und einer wissenschaftlichen Institution – wie gut sind wir? – dienen und Ansporn zur wissenschaftlichen Weiterentwicklung sein.

Die Wissenschaft – wo sonst wäre dazu Kompetenz und Urteilskraft zu finden – entwickelt selbst Instrumente der Leistungsbeurteilung. Wachsende Anforderungen an die Qualitätssicherung in der Wissenschaft, aber auch der Wunsch nach Komplexitätsreduktion zur Herstellung von Vergleichbarkeit stellen die *Methoden* der Messung und der Beurteilung in den Mittelpunkt von Qualitätsdebatten. Dabei ist auch der Gefahr einer Verselbständigung der Methoden, der ‚Messung als Selbstzweck‘ zu wehren; angezeigt ist Qualitätsbewertung mit Augenmaß. Wissenschaft sollte nicht an reinen Produktionsfaktoren, sondern am Erkenntnisfortschritt gemessen werden.

Die international üblichen Instrumente der Leistungsmessung sind durch eine quantitative (basierend auf der Berechnung von Indikatoren) und eine qualitative (basierend auf der Beurteilung durch die Fachkollegenschaft, den *peers*) Zugangsweise charakterisiert. Beide Instrumente, für sich allein genommen, sind mangel- und manipulationsanfällig. Der Königsweg zur ‚Vermessung der Wissenschaft‘ liegt in der Verknüpfung beider Instrumente. Daher ist auch eine kombinierte Vorgangsweise, in der alle wesentlichen Faktoren, unter ihnen Publikationen, Drittmittel, Promotions-

---

<sup>1</sup> D. Kehlmann, *Die Vermessung der Welt*, Reinbek b. Hamburg 2005, 39.

und Habilitationsintensitäten, aber auch Auszeichnungen, Kooperationen und akademische Funktionen, in der *scientific community* berücksichtigt werden, die beste Lösung, wenn es wirklich darum geht, die Qualität von Wissenschaft und Forschung zu beurteilen.

Die Situation in Österreich unterscheidet sich nach Vorgehen und herrschender Beurteilungskultur nicht wesentlich von anderen Wissenschaftssystemen. Insofern läuft auch eine Empfehlung zur Förderung und Sicherung der Forschungsqualität sowie zum Umgang mit den entsprechenden Instrumenten auf ein quantitative mit qualitativen Gesichtspunkten verbindendes Verfahren, ein *informed peer review*-Verfahren hinaus. Es sollte das Standardverfahren darstellen, und zwar auf zwei Ebenen:

1. Auf der Ebene der Bewertung der Forschungsleistung österreichischer Forschungseinrichtungen im Vergleich, verortet in den Leistungsvereinbarungen (für Universitäten, für außeruniversitäre Forschungseinrichtungen und für den forschenden Fachhochschulbereich).
2. Auf der Ebene interner Evaluierungen als Grundlage strategischer Forschungsentscheidungen, z.B. im Kontext einer gesuchten Profil- und Schwerpunktsetzung und bei der inneruniversitären Mittelallokation.

Der Wissenschaftsrat empfiehlt für die Durchführung von *informed peer review*-Verfahren, die standardisiert, jedoch für unterschiedliche Fächer- und Disziplinengruppen unterschiedlich ausgelegt sein müssen, die Entwicklung von ‚Drehbüchern‘ unter Beteiligung von zuständigem Ministerium, forschender Einrichtung und Disziplin bzw. Fachrichtung. Empfohlen wird, hier zunächst mit einem Pilotprojekt, etwa für eine naturwissenschaftliche und eine geisteswissenschaftliche Fächer- oder Disziplinengruppe, zu beginnen.

Kurzgefasst wird für die österreichische Situation, unter Einbeziehung weiterer Gesichtspunkte, empfohlen:

- Eine enge Verbindung sowohl quantitativer als auch qualitativer Verfahren im Sinne eines *informed peer review*-Verfahrens unter Berücksichtigung fächer- und disziplinspezifischer Unterschiede.

- Beachtung des Unterschieds zwischen individueller und institutioneller Leistungsmessung und Leistungsbeurteilung (Qualitätsbeurteilung).
- Weiter gefasste Zeiträume zwischen der Durchführung von Leistungsmessungen und Leistungsbeurteilungen (Leistungsvereinbarungen, Evaluierungen, Audits, Akkreditierungen etc.).
- Stärkere Berücksichtigung von Konsequenzen durchgeführter Leistungsmessungen und Leistungsbeurteilungen sowohl aus der Betroffenen- als auch aus der Prüfendenperspektive. Ohne Konsequenzen gehen Prüfprozesse ins Leere.
- Entlastung des Gutachtersystems. Das Gutachtersystem stößt schon lange an seine Grenzen; zu viele Gutachter pro Verfahren, zu viele Verfahren.
- Neujustierung des Drittmittelwesens zugunsten der Förderung von Spitzenforschung. Wo Drittmittel der Kompensation sonst fehlender Mittel dienen, verfehlen sie ihren Sinn.
- Akkreditierungshürden höher legen, um eine Nivellierung des wissenschaftlichen und institutionellen Niveaus zu verhindern.
- Wissenschaftliche Grade und entsprechende Anforderungen an die Ausbildung schärfen. Bei akademischen Abschlüssen muss gewährleistet sein, dass sie internationalen Standards entsprechen, Tendenzen zu einem *Dr. light* muss entgegengetreten werden.
- Leistungsvereinbarung und Leistungsbeurteilung als allgemeines Modell. Das Instrument der Leistungsvereinbarung sollte in allen forschenden (und lehrenden) Einrichtungen, die durch öffentliche Mittel finanziert (oder mitfinanziert) werden, Anwendung finden. Dies würde auch der Einheitlichkeit eines in sich differenzier-ten Wissenschaftssystems dienen.

## Einleitung

Fragen der Qualitätsbeurteilung und der Qualitätssicherung spielen in der Wissenschaft eine wichtige Rolle. Das war, wenn sich Wissenschaft auf sich selbst bezieht, immer so, gewinnt aber in der modernen Wissenschaftsentwicklung eine ständig wachsende Bedeutung. Der Grund liegt im Werden der modernen Gesellschaft zur Wissensgesellschaft, in der die Abhängigkeit von den Leistungen der Wissenschaft zunimmt, aber auch im Wachstum des Systems Wissenschaft selbst. Während noch bis ins 20. Jahrhundert hinein wenige herausragende Wissenschaftler\* das Leistungsniveau der Wissenschaft und deren Beurteilungsmaßstäbe bestimmten, in gewisser Weise also übersichtliche Verhältnisse herrschten, ist dieses System heute aufgrund seines ungeheuren Wachstums so unübersichtlich geworden, dass es besonderer Anstrengungen und besonderer Instrumente bedarf, um wissenschaftliche Leistungen und damit auch das System Wissenschaft selbst, in seinen jeweils relevanten Teilen, zu beurteilen und seine Leistungsfähigkeit zu sichern. Das gilt in erster Linie für den Bereich der Forschung, etwa der universitären Forschung, der, dem Humboldt-Ideal einer forschungsnahen Lehre entsprechend, eine zentrale Rolle, auch in institutionellen Dingen, d.h. bei der Beurteilung des Leistungsprofils einer Einrichtung insgesamt, hier einer Universität, zukommt (wobei nicht übersehen werden soll, dass der Ausdruck ‚Forschung‘ mittlerweile inflationär, d.h. auch für Tätigkeiten, die nur irgendwie mit Forschung in Verbindung stehen, verwendet wird). Verfahren der Messung von Qualität in der Forschung dienen insofern sowohl der Beurteilung dieser Leistungen selbst als auch der Beurteilung desjenigen institutionellen Rahmens, in dem sie erbracht werden.

In diesem Prozess spielen quantitative Methoden, in erster Linie bibliometrische Methoden, eine wesentliche Rolle. Hinzu tritt in der institutionellen Beurteilung eine zunehmende Bedeutung des Drittmittelsektors, wiederum quantitativ beurteilt. *Wissenschaftssystematisch* bzw. innerwissenschaftlich geht es um die Sicherung transdisziplinärer, d.h. für alle Disziplinen geltender, und disziplinärer, den besonderen disziplinären Gegebenheiten folgender, Standards, *wissenschaftspolitisch* vor dem Hin-

---

\* Die hier verwendeten personenbezogenen Ausdrücke beziehen sich, wenn nicht anders vermerkt, gleichermaßen auf Frauen und Männer. Ferner zur Terminologie: Die Ausdrücke ‚Evaluation‘ und ‚Evaluierung‘ werden (sofern sie nicht in Zitaten auftreten) so verwendet, dass ‚Evaluation‘ die Institution der Evaluierung bedeutet, ‚Evaluierung‘ entsprechend die Praxis der Evaluation.

tergrund einer Verschiebung der Finanzierungsgewichte in der Forschung um die Transparenz von Mittelallokationen und, auf ihr aufbauend, um ein entsprechendes fachliches wie institutionelles ‚Rating‘.

Während die Notwendigkeit von Maßnahmen der Leistungsmessung, auch mit den erwähnten Methoden, nicht in Frage steht, richtet sich Kritik gegen eine befürchtete Ausschließlichkeit des Einsatzes quantitativer Verfahren über alle Disziplinen hinweg und damit zu Lasten einer qualitativen Beurteilung. So wird darauf hingewiesen, dass bei Anwendung ausschließlich quantitativer Verfahren unterschiedlichen Wissenschaftskulturen – z.B. mit Blick auf die unterschiedlichen Verhältnisse in den Naturwissenschaften und den Geisteswissenschaften – nicht Rechnung getragen wird und hier die Zahl der Publikationen, die Zahl der Promotionen und Habilitationen sowie die Höhe der akquirierten Drittmittel qualitative Aussagen in den Hintergrund treten lassen: „Die Verlagerung hin zu einer quantitativen Erhebung der Forschungsleistung fördert (...) die Entwicklung einer outputorientierten Forschung, die qualitative Ansprüche zwangsläufig zurückstellt und zugleich Gefahr läuft, andere Bereiche des Wissenschaftsbetriebes, wie etwa die Lehre oder das Engagement in der akademischen Selbstverwaltung, zu vernachlässigen.“<sup>2</sup> Ein derartiges Vorgehen, so wird argumentiert, reduziert nicht nur die wissenschaftliche Leistung auf Zählbares, es verändert auch das Anreizsystem. Letztlich zählt nur als Leistung, was messbaren Kriterien folgt. Außerdem droht sich das Beurteilungssystem zu verselbständigen; es wird, oft ausgelagert in ‚Agenturen‘, zu einem System im System.

In dieser Situation kommt es in der Wissenschaft, so auch in der Universität, darauf an, sich einerseits zu einer strengen Form der Qualitätsbeurteilung und der Qualitätssicherung zu bekennen und alle diejenigen Instrumente der Leistungsmessung und Leistungsbeurteilung zu nutzen, die heute zur Verfügung stehen, andererseits einer Entwicklung zu wehren, in der sich diese Instrumente verselbständigen. Für das österreichische Wissenschaftssystem, speziell das Universitätssystem, bedeutet dies, sich der entsprechenden Instrumente zu versichern und in Abstimmung mit dem zuständigen Bundesministerium klugen Gebrauch von ihnen zu machen. ‚Ort‘ einer entsprechenden Abstimmung, wenn es um die Darstellung erbrachter Leistungen und die Bestimmung zukünftiger Leistungen im Verhältnis Universität – Ministe-

---

<sup>2</sup> T. Grapatn u.a., Kriterien zur Messung der Forschungsleistung an der Heinrich-Heine-Universität Düsseldorf. Abschlussbericht der AG Leistungsparameter, Düsseldorf 2012, 19.



rium geht, sind die Leistungsvereinbarungen. In ihnen wird Rechenschaft über die Maßnahmen zur Qualitätssicherung abgelegt und werden institutionell verbindliche Verabredungen über die weitere Entwicklung von Forschung und Lehre getroffen.

In den folgenden Kapiteln 1 und 2 wird zunächst der *status quo* der Mess- und Beurteilungsverfahren erhoben, und zwar in internationaler und in nationaler Perspektive, gefolgt von Kapitel 3, das in einem ersten Teil eine Beurteilung dieser Verfahren enthält und in einem zweiten Teil eine solcherart vorgenommene ‚Metaanalyse‘ auf die spezifische Wissenschaftssituation, speziell Universitätssituation, in Österreich bezieht, verbunden mit entsprechenden Empfehlungen. Insgesamt geht es darum, gegebene Verfahren in der Forschung einer Leistungs- bzw. Qualitätsbeurteilung und einer Sicherung entsprechender Leistungen zu beurteilen und das Maß ihrer Anwendung im österreichischen Wissenschaftssystem, speziell im österreichischen Universitätssystem, zu bestimmen.

## 1. Mess- und Beurteilungsverfahren in internationaler Perspektive<sup>3</sup>

Studien zur Messung der Qualität von Forschung mit Hilfe quantitativer Indikatoren waren vor den 1980er Jahren noch von ausschließlich akademischem Interesse; erst später – im wesentlichen aufgrund der Krise öffentlicher Haushalte – erhielten Studien dieser Art eine hochschul- und wissenschaftspolitische Bedeutung. Die Mittelvergabe für Wissenschaft und Forschung wurde mit den Ergebnissen hauptsächlich quantitativer Bewertungen verbunden. In den 1990er Jahren wurden auch *peer review*-Verfahren, die ursprünglich zur Begutachtung von zur Publikation vorgelegten Arbeiten eingesetzt wurden, zu einem Instrument der Beurteilung wissenschaftlicher Leistungen. Preise, Stipendien und Schlüsselpositionen in Wissenschaft, Forschung und Lehre werden aufgrund des Einsatzes dieser Instrumente vergeben.

*Peer review*-Verfahren zur Bewertung von Qualität sind eng mit der Entwicklung der modernen Wissenschaft verbunden. Ohne dieses formalisierte Verfahren im Sinne eines ‚organisierten Skeptizismus‘<sup>4</sup> hätte die Qualität von Forschung nicht systematisch und dauerhaft sichergestellt werden können. Die Befürworter des *peer review* halten dementsprechend dieses Verfahren zur Prüfung und Legitimation wissenschaftlicher Arbeit für geeigneter als jedes andere Verfahren – es sei wissenschaftsadäquat und trage zur Vereinheitlichung von Bewertungskriterien und formalen Anforderungen bei. Die Stärke des *peer review*-Verfahrens liegt tatsächlich in der Bewertung von Publikationen, Forschungsanträgen und Wissenschaftlern sowie kleineren Forschungsgruppen. Auch kann (und sollte) es vor allem dort eingesetzt werden, wo quantitative Instrumente der Forschungsbewertung an ihre Grenzen stoßen.

Kritiker des *peer review*-Verfahrens sehen vor allem drei Schwächen: (1) Unterschiedliche Gutachter stimmen kaum in der Bewertung ein und derselben wissenschaftlichen Arbeit überein (fehlende Reliabilität); (2) gutachterliche Empfehlungen weisen systematische Urteilsfehler auf, die auch auf nicht-wissenschaftlichen Kriterien, z.B. Geschlecht oder Nationalität, basieren (fehlende Fairness); (3) der Zusammenhang zwischen dem Urteil im *peer review* und der Qualität der begutachteten Arbeit ist gering (fehlende Validität).

---

<sup>3</sup> Dieses Kapitel beruht auf einem Gutachten, vorgelegt von Dr. Lutz Bornmann, Max-Planck-Gesellschaft (MPG) München, im Auftrag des Österreichischen Wissenschaftsrates (2013).

<sup>4</sup> R. K. Merton, *The Sociology of Science. Theoretical and Empirical Investigations*, Chicago 1973.

Seit Ende der 1980er Jahre gewinnt die *ex post*-Bewertung von Forschungsleistungen durch bibliometrische Indikatoren zunehmend an Bedeutung. Bei dieser Form der Bewertung werden Publikations- und Zitierhäufigkeiten als numerisches Relativ für die Güte und Wirksamkeit einer wissenschaftlichen Arbeit verwendet und als Entlastung des Gutachterwesens verstanden. Zitat und Verweis zeigen an, dass wissenschaftliche Arbeiten als relevant für den Wissenszuwachs in der wissenschaftlichen Gemeinschaft, aber auch als relevant für die Gesellschaft eingeschätzt werden. Der Grund für die Popularität, die bibliometrische Indikatoren genießen, wird in der Skepsis der Wissenschaftspolitik und der Wissenschaftsverwaltung gegenüber den wenig verallgemeinerbaren Ergebnissen von *peer review*-Verfahren gesehen. Von bibliometrischen Indikatoren könne man nicht nur unabhängiger und objektivere Bewertungen als von Fachgutachtern erwarten, sie seien auch billiger und weit weniger zeitaufwendig.

Während Publikationshäufigkeit als Indikator wissenschaftlicher Produktivität weitgehend anerkannt ist, werden gegen die Verwendung von Zitierhäufigkeiten als Indikator für die Qualität einer wissenschaftlichen Arbeit seit Jahren zahlreiche Bedenken geäußert: (1) Wissenschaftliche Qualität sei ein komplexes Phänomen und könne nicht über ein eindimensionales Messkonstrukt wie die Zitierhäufigkeit gemessen werden; (2) die Zitierweise sei von Eigeninteressen beeinflusst; (3) relevante Publikationen könnten übersehen werden; (4) Zitationshäufigkeiten seien fälschungsanfällig etc.. Da Studien belegen, dass Zitierungen mit dem Urteil von Fachgutachtern substantiell korrelieren, wird ihre Anwendung in der Forschungsevaluation jedoch generell als gerechtfertigt angesehen; bibliometrische Datenbanken wie *Web of Science* oder *Scopus* stellen hier eine wichtige Grundlage dar. In der Forschungsevaluation werden diese Datenbanken für die Bewertung von wissenschaftlicher Arbeit mittels (mehr oder weniger) standardisierter Verfahren ausgewertet. Einer dieser bibliometrischen Indikatoren ist der von J. E. Hirsch<sup>5</sup> entwickelte *h-Index*, mit dem bisherige Leistungen und die Identifikation erfolgreicher Wissenschaftler in einem Fachgebiet mit einer einzigen Zahl gemessen werden könnten. Mittlerweile werden dem *h-Index* allerdings erhebliche Schwächen attestiert und Varianten entwickelt. Eine ähnliche Prominenz wie der *h-Index* besitzt der *Journal Impact Factor (JIF)*, der die Zitierung

---

<sup>5</sup> J. E. Hirsch, An Index to Quantify an Individual's Scientific Research Output, Proceedings of the National Academy of Sciences of the United States of America 102 (2005), No. 46, 16569-16572.

gen in einer wissenschaftlichen Zeitschrift in einem bestimmten Zeitraum ermittelt und nicht nur zur Bewertung der Zeitschrift selbst, sondern auch zur Bewertung einzelner Publikationen sowie zur Evaluierung von Forschern oder Forschungsinstituten herangezogen wird.

Weder *h-Index* noch *JIF* sehen eine Normierung vor. Es stellt sich daher die Frage, wie besonders hoher oder niedriger *output* im Vergleich messbar ist. Die Bibliometrie normiert zum einen Zitierungen im Hinblick auf das Publikationsjahr; ältere Publikationen sollten öfter zitiert worden sein als jüngere. Zum anderen ist es erforderlich, Zitierungen im Hinblick auf ein Fachgebiet zu normieren. Um diese Zeit- und Fachnormierung für die Wirkung einer bestimmten Publikation herzustellen, wird üblicherweise ein referentielles Publikationsset zusammengestellt, anhand dessen es möglich sein soll, zu beurteilen, welche Wirkung die betreffende Publikation im Vergleich zu ähnlichen in einem bestimmten Zeitraum erzielt hat. Zu berücksichtigen ist, dass sich diese Standards in der Regel auf Zeitschriften beziehen, entweder auf einzelne Zeitschriften (zeitschriftenbasierter Indikator, ZBI) oder auf für ein Fachgebiet relevante Zeitschriftensets (fachgebietsbasierter Indikator, FBI). Beide Indikatoren kommen in der Bibliometrie zur Anwendung. Dabei spricht vieles für eine Präferenz fachgebietsbasierter Indikatoren als Grundlage einer Normierung: Bei ausschließlicher Verwendung des ZBI besteht die Gefahr, dass aus pragmatischen Gründen Zeitschriften mit geringem Renommee für Publikationen genutzt werden; das Publikationsjahr als Referenzwert hat *per se* nichts mit der Qualität einer Publikation zu tun; unterschiedliche Zitierraten in Fachgebieten sagen nichts über die Qualität, sondern ausschließlich etwas über das Zitierverhalten in einem Fachgebiet aus. Die Bibliometrie präferiert aus diesen Gründen als Referenzrahmen die Verwendung der Zitiergewohnheiten eines ganzen Fachgebietes statt die einer einzelnen Zeitschrift. Auch hier gibt es Optimierungsvorschläge bei der Indikatorenbildung, die sich vor allem auf die Verwendung der fachspezifischen Referenzsets und die Vermeidung des arithmetischen Mittelwertes durch den Gebrauch von Perzentilen beziehen.

Im Bereich der Geistes- und Sozialwissenschaften lassen sich bibliometrische Verfahren nur begrenzt einsetzen. Die gängigen Literaturdatenbanken wie *Web of Science* oder *Scopus* erfassen die Publikationen der Geistes- und Sozialwissenschaften nicht, d.h. Arbeiten, die weniger in Form von Beiträgen in Zeitschriften als vielmehr in Form von Monographien oder Beiträgen in Sammelbänden, oft auch in

eher nicht-wissenschaftlichen Medien erscheinen. Darüber hinaus ist der Erkenntnisfortschritt in den Sozial- und Geisteswissenschaften häufig durch einen längeren Zeithorizont geprägt. Späte Wirkungen wissenschaftlicher Arbeiten sind hier tendenziell häufiger zu finden, werden aber in gängigen Literaturlauswertungen nicht berücksichtigt. Daher stehen andere Indikatoren, mit denen Forschungsleistungen gemessen werden können und die auch für die Messung von wissenschaftlichen Leistungen in den Naturwissenschaften gelten, im Vordergrund. Zu diesen gehören:

- Eingeworbene Drittmittel und ihre Höhe (national wie international)
- Eingeladene Teilnahme an nationalen oder internationalen wissenschaftlichen Kongressen (z.B. Hauptvortrag oder Organisation einer Arbeitsgruppe)
- Gemeinsame Forschungsprojekte (z.B. Gründung eines Forschungsschwerpunktes oder Betreiben eines Forschungszentrums)
- Berufungen und Auszeichnungen (z.B. Wahl in Akademien, Ehrendoktorwürden)
- Gastprofessuren (im In- oder Ausland)
- Vertretung von Professuren
- Fremdfinanzierte Forschungs- und Lehraufenthalte im Ausland
- Gutachtertätigkeiten
- Ämter in der akademischen Selbstverwaltung
- Herausgehobene Funktionen in der *scientific community*

Der Nachteil bei der Verwendung dieser Indikatoren besteht darin, dass sie keine Normierung erlauben, wie sie in bibliometrischen Verfahren zumindest möglich sind. Normierungen aber sind eine wesentliche Voraussetzung für einen objektiven Zugang zur Leistungsbeurteilung. Bemühungen sind im Gange, in diesem Sinne auch für die Geistes- und Sozialwissenschaften anwendbare Qualitätskriterien zu entwickeln; dabei wird unter anderem mit fachspezifischen Datenbanken als Grundlage bibliometrischer Messmethoden experimentiert.

Seit Mitte der 1990er Jahre lässt sich ein Trend zur Intensivierung von *Evaluationen* in der Wissenschaft im allgemeinen und in der Forschung im besonderen beobachten, die auch den sozialen, kulturellen und ökonomischen, in manchen Fällen auch den ökologischen Nutzen der Forschung berücksichtigen. Bekanntestes Beispiel ei-

ner derartigen Evaluationspraxis ist das *Research Assessment Exercise* in Großbritannien; dieses wurde in den vergangenen Jahren in ein *Research Excellence Framework* (REF) überführt. In ihm wird die Forschungsleistung qualitativ beurteilt und quantitativ mit bibliometrischen Methoden unterlegt. Zusätzlich wird die gesellschaftliche und wirtschaftliche Wirksamkeit der Forschung in allen Fachbereichen auf dem Wege von Fallstudien durch Fachgutachter geprüft. Die Entwicklung robuster Methoden zur Bemessung der Wirkung und des Nutzens von Forschung steckt zwar noch in den Anfängen, doch besitzt sie bereits budgetrelevante Wirkungen; die Allokation von Forschungsmitteln orientiert sich immer stärker am (vermuteten) gesellschaftlichen Nutzen von Forschung.

In den Naturwissenschaften haben sich bibliometrische Verfahren als Standard in der quantitativen Forschungsevaluation etabliert, während in den Geistes- und Sozialwissenschaften nach wie vor unklar ist, welche Rolle derartige Verfahren vor dem Hintergrund eines primär geübten *peer review*-Verfahrens spielen können.

Bei der Gegenüberstellung von *peer review* und bibliometrischen Verfahren wird häufig übersehen, dass beide Instrumente eng miteinander verknüpft sind.<sup>6</sup> Bibliometrische Indikatoren resultieren aus einem formalen wissenschaftlichen Kommunikationsprozess, der sich vor allem auf Veröffentlichungen in Fachzeitschriften stützt und durchaus Elemente eines *peer review*-Verfahrens einschließt. Während einige das qualitative *peer review*-Verfahren und die quantitativen Methoden der Bibliometrie in Konkurrenz zueinander sehen<sup>7</sup>, empfiehlt die Mehrheit, „dass sie sich wechselseitig ergänzen sollten, um die Schwächen des jeweils anderen Verfahrens zu kompensieren“<sup>8</sup>. Erst durch die Interpretation von Experten ließe sich aus den Ergebnissen bibliometrischer Analysen eine qualitative Einschätzung wissenschaftlicher Leistungen entwickeln. Dabei bedürfe ein indikatorengestütztes Verfahren nicht nur der fachkundigen Kommentierung, sondern auch der Korrektur und Ergänzung durch ein aussagefähiges Kontextwissen.

---

<sup>6</sup> Vgl. S. Hornbostel, *Wissenschaftsindikatoren. Bewertungen in der Wissenschaft*, Opladen 1997.

<sup>7</sup> Vgl. B. Bozeman/J. S. Dietz/M. Gaughan, *Scientific and Technical Human Capital. An Alternative Model for Research Evaluation*, *International Journal of Technology Management* 22 (2001), No. 7-8, 716-740.

<sup>8</sup> H.-D. Daniel, *Wissenschaftsevaluation. Neuere Entwicklungen und heutiger Stand der Forschungs- und Hochschulevaluation in ausgewählten Ländern*, Center for Science and Technology Studies (CEST), Bern 2001, 11.

Alles in allem gibt es gute Gründe, das qualitative *peer review* und die quantitativen Methoden der Bibliometrie nicht als konkurrierende, sondern als sich ergänzende Verfahren in der Forschungsevaluation anzusehen. Problematisch ist in diesem Zusammenhang allerdings die Messung der gesellschaftlichen Wirkung von Forschung. Bis hier überzeugende Lösungen gefunden sind, sollte wohl verstärkt auf das *peer review*-Verfahren zurückgegriffen werden.

## 2. Mess- und Beurteilungsverfahren in nationaler Perspektive

### 2.1 Verfahren an österreichischen Universitäten<sup>9</sup>

In Österreich ist die Messung und Beurteilung von Forschungsleistungen vor dem Hintergrund der Rahmenbedingungen, die der Gesetzgeber mit dem UG 2002 bestimmt hat und das für Wissenschaft und Forschung zuständige Ministerium durch seine Hochschulpolitik verfolgt, zu sehen. Demnach sollen folgende Instrumente zur Beurteilung und Steuerung von Forschung an österreichischen öffentlichen Universitäten eingesetzt werden: Leistungsvereinbarungen, Zielvereinbarungen, Hochschulraumstrukturmittel (vormals Formelbudget), Wissensbilanzen, Evaluierungen, Qualitätsmanagement. Die Leistungsvereinbarung stellt hier das zentrale Steuerungsinstrument für die Finanzierung der österreichischen Universitäten, neuerdings auch der Österreichischen Akademie der Wissenschaften (ÖAW), durch die öffentliche Hand dar.

Die wichtigsten Ziele im Bereich Forschung für die dritte universitäre Leistungsvereinbarungsperiode 2013 bis 2015 sind die Fortführung der Schwerpunkt- und Profilbildung, die Intensivierung von Kooperationen, der Ausbau der Internationalisierung, der strategische Ausbau der Forschungsinfrastruktur und die weitere (quantitative und qualitative) Entwicklung der Drittmittelinwerbung. Die Umsetzung der Leistungsvereinbarung spiegelt sich in den universitätsintern definierten Zielvereinbarungen mit den Organisationseinheiten wider. Im Bereich Forschung werden bei einigen Universitäten verstärkt Indikatoren genutzt, die die Zielerreichung und den Erfolg der Umsetzung von Vorhaben messbar machen sollen. Die mit dem *bmwf*, seit 2014 *bmfwf*, ausverhandelten Zielwerte sind – vor dem Hintergrund stagnierender Budgets – mehrheitlich konservativ angesetzt.

Bei den Zielvereinbarungen gehen die Universitäten<sup>10</sup> nach einem relativ standardisierten und strukturierten Prozess vor, bei dem mit Hilfe von Arbeitsbehelfen, Vorla-

---

<sup>9</sup> Dieses Kapitel beruht zu großen Teilen (Abschnitt 2.1) auf einem Gutachten, vorgelegt von Dr. Karl-Heinz Leitner, Austrian Institute of Technology, durchgeführt im Auftrag des Österreichischen Wissenschaftsrates (2013).

<sup>10</sup> Die folgende Darstellung basiert auf der Analyse von Unterlagen der Universitäten (Leistungsvereinbarungen, Zielvereinbarungen, Satzungen), auf Ergebnissen früherer Studien und Interviews mit Vertretern der Universitäten, wobei vor allem Mitarbeiter aus den Rektoraten, die für Zielvereinbarung, Forschungsdokumentation, Qualitätsmanagement und Evaluation zuständig sind, befragt wurden. Ausgewählt für eine vertiefende Analyse, die den Schlussfolgerungen zugrundeliegt, wur-



gen, Fragenkatalogen und Indikatoren Forschungsleistungen intern erfasst und bewertet werden. Der Gesetzgeber hat hier keine methodischen Vorgaben gemacht, doch lässt sich insgesamt in den vergangenen Jahren ein Trend zur Quantifizierung und Objektivierung konstatieren. Die Ausgestaltung der Zielvereinbarungen und die zunehmend systematische Nutzung von Indikatoren dokumentieren diese Entwicklung. Dennoch gibt es Unterschiede zwischen den Universitäten; das Ausmaß der Quantifizierung und die Nutzung von Evaluierungsergebnissen hängen stark von der historisch gewachsenen Universitätskultur und den Vorgaben der Universitätsleitung ab.

Im Rahmen einer Novelle des UG wurde im Sommer 2012 das Instrument Hochschulraum-Strukturmittel (HRSM) eingeführt. Neben dem Grundbudget, das weiterhin auf Basis von Leistungsvereinbarungen definiert wird, wird für die LV-Periode 2013 bis 2015 erstmals die leistungsorientierte Finanzierung in Form des Formelbudgets als Teil des Globalbudgets durch einen neuen Allokationsmechanismus abgelöst. Mit der Anzahl der prüfungsaktiven Studierenden, der Anzahl der Studienabschlüsse und den Einnahmen aus F&E-Projekten wurden beim Formelbudget drei Indikatoren verwendet, die in adaptierter Form nun beim Verteilungsmodell der Hochschulraum-Strukturmittel<sup>11</sup> zur Anwendung kommen.

Mit dem UG 2002 wurden österreichische Universitäten verpflichtet, Wissensbilanzen zu erstellen. Dieses Instrument, orientiert am Paradigma des *New Public Management*, wird seit seiner Einführung kontrovers diskutiert und vielfach kritisiert; es steht für eine ausschließlich quantitative Beurteilung von Forschungsleistungen. Hier sind mehr als 30 Indikatoren (durch eine entsprechende Verordnung genauer definiert), die über die genutzten Ressourcen, Prozesse der Weiterentwicklung und Leistungen der Universität Auskunft geben, jährlich auszuweisen. Zentrale Kenngrößen zur Anzahl des Personals, zur Anzahl von Berufungen, zur durchschnittlichen Studiendau-

---

den die Universitäten Graz, Innsbruck, Salzburg und Wien (diese haben jeweils auch geisteswissenschaftliche Fakultäten), die TU Graz als Vertreterin einer Technischen Universität, die Medizinische Universität Innsbruck als Repräsentantin der Medizinischen Universitäten sowie die Universität für Musik und darstellende Kunst Graz, stellvertretend für die Kunstuniversitäten.

<sup>11</sup> Die Hochschulraumstrukturmittel sind Teil der Hochschulmilliarde und mit 450 Mio. Euro dotiert. Ein Teilbetrag von 63 Mio. Euro wurde im September 2013 für die Anschubfinanzierung von Kooperationen nach Ausschreibung vergeben. Antragsberechtigt waren die öffentlichen Universitäten. Voraussetzung für die Vergabe der Projektmittel war die Beteiligung mindestens einer weiteren Institution aus den Bereichen Wissenschaft, Hochschulen, Kunst und Kultur oder der Wirtschaft. Weitere Kooperationsausschreibungen sind geplant (vgl. Pressemeldung des *bmwf* vom 17.9.2013).

er, zur Mobilität, zu Drittmiteleinahmen und zur Anzahl der Veröffentlichungen werden von allen öffentlichen österreichischen Universitäten erfasst. Die Aussagekraft der dafür genutzten Kennzahlen wird mehrheitlich als gut beurteilt. Schwächen hinsichtlich der Erfassung der Qualität des Forschungsoutputs suchen die Universitäten unter anderem durch eine Gewichtung unterschiedlicher Publikationstypen, durch Herausfiltern zusätzlicher Kennzahlen (z.B. in Form der Differenzierung zwischen nationalen und internationalen Vorträgen) und durch die Nutzung von Rankings zu beheben. Auch die Durchführung von regelmäßigen Evaluierungen ist für die Universitäten nach Maßgabe ihrer Satzung verpflichtend, doch nicht im Detail vorgeschrieben. Die Praxis der Universitäten orientiert sich hier in der Regel an internationalen Standards. Desweiteren sind die Universitäten gemäß Universitätsgesetz verpflichtet, ein Qualitätsmanagementsystem einzuführen; nach dem Qualitätssicherungsgesetz von 2011 sind sie ferner verpflichtet, dieses Qualitätsmanagementsystem im Siebenjahresrhythmus zertifizieren zu lassen. Dabei steht es ihnen frei, das Audit durch die österreichische Qualitätssicherungsagentur AQ Austria oder durch einen ausländischen Anbieter durchführen zu lassen.<sup>12</sup> Entsprechende Vorhaben (unterschiedlichste quantitative und qualitative Verfahren) werden im Rahmen der Leistungsvereinbarungen hinsichtlich Einsatz und Effekt dargestellt.

Rankings und Benchmarking spielen in Österreich – abgesehen von einigen Pilotprojekten und Überlegungen, sich dabei am Projekt *U-Multi-Rank*<sup>13</sup> der Europäischen Union zu beteiligen – bislang eine untergeordnete Rolle.

Die Universitäten reagieren auf nationale Anforderungen und Strategien (Leistungsvereinbarungen durch das *bmwf/bmwf*) wie auch auf den internationalen Trend zur Messung und Beurteilung von Forschungsleistungen (basierend auf Rankings, Indikatoren zur Wissenschaftskultur, der Existenz von Kriterien für erfolgreiche Wissenschaftskarrieren, der Bedeutung von Drittmitteln etc.); allen diesen Perspektiven ist

---

<sup>12</sup> Die Universität Graz unterzog 2013 als erste heimische Universität ihr internes Qualitätssicherungssystem einem umfassenden Audit durch die finnische Agentur Finheec (Finnish Higher Education Evaluation Council), mit positivem Ergebnis. Sie erhält für die Einhaltung internationaler Qualitätssicherungskriterien in Forschung, Lehre und gesellschaftlicher Wirkung das Gütesiegel der Agentur für die kommenden sieben Jahre ohne Auflagen (vgl. <http://diepresse.com/home/bildung/universitaet/1453954/Guetesiegel-fuer-Uni-Graz?from=suche.intern.portal> (Stand 20.9.2013).

<sup>13</sup> <http://www.multirank.eu/> (Stand 22.7.2013).

eine stärkere Quantifizierung inhärent. Dies spiegelt sich auch in den Leistungs- und Zielvereinbarungen wider.

Die Stabstellen bzw. Servicestellen der Universitäten bereiten in der Regel Informationen und Kennzahlen zur Erstellung von Leistungsvereinbarungen, Zielvereinbarungen, Wissensbilanzen und Evaluierungen auf. Des Weiteren werden vielfach – z.B. als Berechnungsgrundlage einer leistungsorientierten Mittelvergabe (LOM) oder auf Wunsch der Fakultäten – anspruchsvollere bibliometrische Analysen durchgeführt (Zitationsanalysen oder Berechnungen des *h-Index*).

Einigkeit besteht unter den befragten Universitäten, dass ein direkter Vergleich des Publikationsoutputs zwischen unterschiedlichen Disziplinen nicht möglich und auch nicht sinnvoll ist, etwa der zwischen der Physik und der Soziologie. Das gilt selbst im Rahmen spezialisierter Universitäten wie den Technischen und den Medizinischen Universitäten. Anders im Falle eines Vergleichs zwischen Universitäten mit etwa gleichem Disziplinspektrum; dieser lässt durchaus differenzierte Aufschlüsse über die Leistungsfähigkeit von Universitäten zu.

Die Universitäten stehen intern vor der Herausforderung, unterschiedliche Forschungseinheiten und Forschungsschwerpunkte zu bewerten und Mittel leistungsbezogen nach bestimmten Kriterien zu verteilen. Das Credo der befragten Universitäten lautet hier, Fakultäten, Institute und andere Forschungsbereiche so zu behandeln, wie es deren Wissenschaftskultur entspricht (also kontextbezogen). Dabei werden unterschiedliche, meist kombinierte Strategien gewählt:

- Selbstreports (bei Evaluierungen). Sie bieten den Spielraum, Spezifika und Besonderheiten darzustellen, die dann durch *peers* gewertet werden.
- Auswahl der 5 bis 15 besten wissenschaftlichen Publikationen. Sie ermöglicht es, ein besonderes Forschungsprofil hervorzuheben.
- Eine disziplinenmäßig abgestimmte Verwendung von Rankings unter wissenschaftlichen Publikationstypen. Diese werden entweder von den Organisationseinheiten selbst definiert oder von internationalen Gremien übernommen (z.B. *European Reference Index for the Humanities*, *European Science Foundation*).

Vor allem die Bewertung von in sich heterogenen Forschungsfeldern stellt große methodische Herausforderungen. Das wird schon in dem Versuch erkennbar, die fünf wichtigsten Publikationen herauszufiltern. Ein *peer review*-basiertes Evaluationsverfahren würde hier eine große Zahl von *peers* erfordern, d.h. ein sehr aufwendiges Verfahren. In vielen Fällen wird daher auf eine quantitativ dargestellte Beurteilung von Publikationen auf der Ebene eines übergeordneten Forschungsfeldes verzichtet (wie zu einem guten Teil auch im Rahmen der Leistungsvereinbarungen auf der Ebene der Gesamtuniversität). Bei relativ homogenen Forschungsprofilen (z.B. TU Graz oder Medizinische Universität Innsbruck) werden ausgewählte Publikationstypen (zumeist *Web of Science*-Publikationen) pro Forschungsschwerpunkt ausgewiesen. Einen Weg, auch relativ breit ausgerichtete Forschungsfelder quantitativ zu bewerten, hat die Universität Innsbruck beschritten, indem unterschiedliche Publikationsformen gewichtet und diese für den gesamten Forschungsschwerpunkt aggregiert dargestellt werden.

Evaluierungen werden als besonders aussagekräftiges Instrument des Qualitätsmanagements und damit auch der strategischen Universitätsentwicklung gesehen; ihre Ergebnisse fließen in die Zielvereinbarungen ein. Die derzeitige Praxis – eine alle fünf Jahre stattfindende Evaluierung – wird von den befragten Universitäten als angemessen betrachtet. In der Regel werden Leistungsvereinbarung, Zielvereinbarung und Evaluierung zeitlich aufeinander abgestimmt. Eine Umsetzung von Evaluierungsergebnissen wird von Workshops begleitet.

Auch bei der Bewertung der geisteswissenschaftlichen und der künstlerischen Forschung wird der Weg einer quantitativen Bewertung eingeschlagen. Hier stehen, z.B. im Vorgehen der Kunstuniversität Graz, quantitative Verfahren anhand von Kennzahlen im Mittelpunkt, wobei die Beurteilung der Qualität künstlerisch-wissenschaftlicher Forschungsleistungen nur Druckerzeugnisse berücksichtigt, insbesondere die Monographie, nicht die künstlerische Produktion.<sup>14</sup> Im englischsprachigen Bereich hat sich dagegen die Wertschätzung von der Monographie auf *peer-reviewte* Artikel in Fachzeitschriften verlagert.

---

<sup>14</sup> Das Programm PEEK des FWF ist unter anderem durch den Versuch charakterisiert, eine differenzierte qualitative Beurteilung künstlerisch-wissenschaftlicher Leistungen vorzunehmen.

Neben der Nutzung intern erstellter oder externer Rankinglisten (z.B. *European Reference Index for the Humanities*, ERIH) werden etwa an der Universität Graz die Reputation und die Transferleistung unter Verwendung von Kennzahlen diskutiert (z.B. Tätigkeiten in Kommissionen, Funktionen in Zeitschriften, Gutachtertätigkeit, Habilitationsverfahren, Vorträge), die besser über einzelne Fächer hinweg verglichen werden können. Bewertungsverfahren, die den *impact* der Forschung zur Lösung gesellschaftlicher Probleme erfassen und häufig im Zusammenhang mit der Bewertung der Forschungsleistungen in den Geisteswissenschaften diskutiert werden, werden bislang kaum genutzt.

Im allgemeinen existiert bei der Bestimmung des Budgets kein direkter Zusammenhang zwischen Zielerreichung und Finanzierung. Gleichwohl schwächt die Nichterreichung von Zielen mittelfristig die Verhandlungsposition und erschwert die Finanzierung von längerfristigen Vorhaben und Projekten. Die Ergebnisse von Evaluierungen haben in der Regel Einfluss auf die universitätsinterne Verteilung von Mitteln. Was die Nutzung von Leistungs- oder sogenannten ‚Outputindikatoren‘ betrifft, so werden diese bei Evaluierungen als komplementäre Informationsgrundlage betrachtet und klassischerweise als *informed peer review* bezeichnet. Bei den Zielvereinbarungen stellen diese Indikatoren eine wichtige Orientierungsgröße dar; sie sollen (indirekt) Anreize zur Zielerreichung schaffen.

Von Seiten des Universitätsmanagements wird der größte Vorteil der stärkeren Nutzung von Kennzahlen in der Objektivierung von Entscheidungen gesehen. Rankings und Benchmarking spielen noch eine untergeordnete Rolle. Es besteht jedoch die Absicht, auch diese Instrumente mittelfristig stärker zu nutzen. Hier wird von den Universitäten betont, derartige Vergleiche und Analysen auf der Ebene der Fachbereiche zu installieren.

Profil- und Schwerpunktbildung in der Forschung stellen wichtige hochschulpolitische Ziele in Österreich dar. Forschungsschwerpunkte werden im Rahmen der Leistungsvereinbarungen hervorgehoben, doch fehlt es noch an einer nachvollziehbaren Bewertung im Rahmen der Leistungsvereinbarungen. Gesamtuniversitäre Forschungsschwerpunkte werden von Universitäten meist als noch zu jung beurteilt (sie bestehen seit rund fünf bis acht Jahren), um sie schon sinnvoll evaluieren zu können. Dagegen gewinnen in der Beurteilung von Forschungsexzellenz, individueller wie insti-

tutioneller Exzellenz, hochrangige Förderungen und Preise, etwa durch den *European Research Council* (ERC) zunehmend an Bedeutung.

## 2.2 Verfahren an außeruniversitären Einrichtungen

### 2.2.1 Beispiel *Fonds zur Förderung der wissenschaftlichen Forschung* (FWF)

Der FWF als wichtigste Institution zur Förderung der Grundlagenforschung in Österreich verwendet zur Begutachtung und Evaluation der von ihm finanzierten Projekte ausschließlich das *peer review*-Verfahren. Pro Jahr müssen für die Beurteilung beantragter Forschungsprojekte um die 5.000 Gutachten verfasst werden. Bei der Auswahl der Gutachter wird streng darauf geachtet, die wichtigsten internationalen Vertreter der Disziplin zu gewinnen; diese müssen mindestens das gleiche Qualifikationsniveau wie die Antragssteller aufweisen. Die Rekrutierung der Gutachter gestaltet sich durch die „schleichende Systemüberlastung, ausgelöst durch eine exzessive Evaluationspraxis“ immer schwieriger<sup>15</sup>. Die Begutachtung oft sehr unterschiedlicher Exzellenzinitiativen, umfangreiche Begutachtungen, die bereits bei Dissertationskonzepten verlangt werden, die wachsende Begutachtungstendenz auch bei der Förderung selbst von Reise-, Konferenz- und Workshopkosten, die Begutachtung interner Forschungsfreistellungen sowie Zwischenevaluierungen in immer kürzeren Zeitabständen haben dazu geführt, dass 2012 die Zusagequote bei der Nominierung von Gutachtern von 45 auf 33 Prozent zurückgegangen ist; diese werde, so der FWF, weiter sinken.<sup>16</sup>

---

<sup>15</sup> C. Kratky, Präsident des FWF, im Workshop des Wissenschaftsrates zur „Messung und Bewertung von Forschungsleistungen“ am 13.4.2013.

<sup>16</sup> Man könne zunehmend froh sein, Gutachter aus der zweiten und dritten Reihe zu gewinnen; ders., a.a.O..

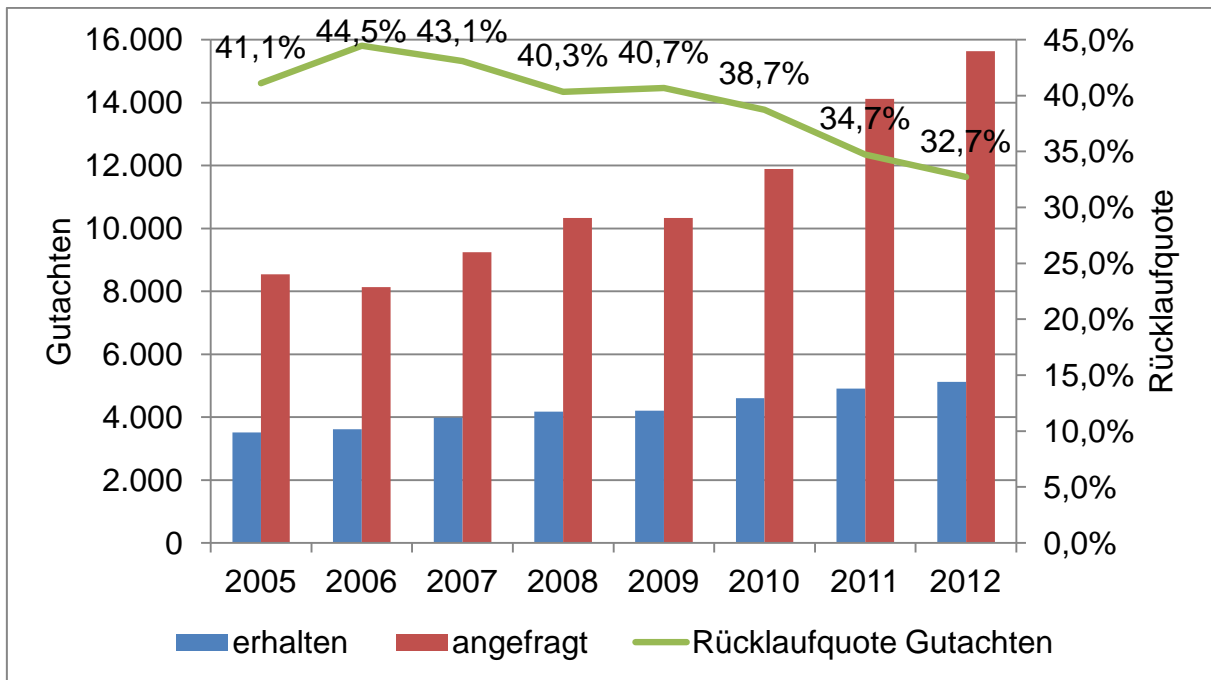


Abb. 1: Rücklaufquoten der Gutachten 2005-2012, C. Kratky 2013.

Quantitative Messmethoden wie *Journal Impact*-Faktoren werden seitens des FWF als zu wenig transparent in der methodischen Generierung und daher kritisch gesehen; dabei werde der Zeitschrift, in der ein Wissenschaftler bzw. ein Autorenteam publiziert, mehr Bedeutung beigemessen als dem Inhalt des Artikels selbst. Der FWF hat, ebenso wie die Herausgeber führender wissenschaftlicher Zeitschriften, die *San Francisco Declaration on Research Assessment*<sup>17</sup>; darin wird unter anderem gefordert, *Journal Impact*-Faktoren nicht als Indikatoren zur Messung von Forschungsleistung zuzulassen, sondern die qualitative Beurteilung in den Vordergrund zu stellen. Der FWF orientiert sich in seiner Begutachtungspraxis ferner an der Methodik der Forschungsevaluation durch das britische *Research Excellence Framework* (REF). Dieses setzt sich aus Verfahren der permanenten und der periodischen Forschungsevaluation zusammen; das erste Verfahren bezieht sich auf die antragsbasierte Vergabe der öffentlichen Forschungsmittel über die *Research Councils UK* und anderer Drittmittelgeber sowie die Anzahl der Doktoranden und erfolgreichen Promotionsverfahren, das zweite Verfahren auf eine disziplinspezifische Expertengewichtung von Forschungsleistung und der Wirkung (*output* und *impact*) aus-

<sup>17</sup> <http://am.ascb.org/dora/files/SFDeclarationFINAL.pdf> (Stand 14.7.2013).

gewählter Forschungsergebnisse sowie der institutionellen Forschungsinfrastruktur (*environment*).

Die Ergebnisse von *peer review*-Verfahren sind Basis für alle Förderentscheidungen des FWF. Dabei sieht sich der FWF der Förderung aller Wissenschaften in gleicher Weise verpflichtet; eine Quotenregelung zur Steuerung der Mittelverteilung zwischen den einzelnen Fachgebieten gibt es nicht.

#### *Ex ante-Evaluierung (Förderanträge und Förderentscheidung)*

Als Gutachter werden ausschließlich international ausgewiesene ausländische Wissenschaftler eingesetzt. Diese werden vom Präsidium des FWF fachspezifisch ausgewählt; die Anzahl der Gutachter ist von der Höhe der Antragssumme und der Programmkategorie abhängig. Die meisten Gutachten werden auf schriftlichem Wege eingeholt; gelegentlich kommen aber auch Panelreviews zum Einsatz (z.B. bei Schwerpunktprojekten oder bei Exzellenzprogrammen).<sup>18</sup> Es gilt das Verbot der Doppelförderung. Die Beurteilung der Anträge erfolgt laut FWF „nach international anerkannten Qualitätskriterien und nach der Bedeutung des Forschungsprojekts für den Erkenntnisgewinn und die Erweiterung sowie Vertiefung der wissenschaftlichen bzw. künstlerischen oder künstlerisch-wissenschaftlichen Kenntnisse“<sup>19</sup>.

Zur Bewertung von Programmen wie PEEK<sup>20</sup>, KLIF<sup>21</sup>, von START-Programm und Wittgensteinpreis werden eigene Jurys bzw. vergleichbare Gremien (z.B. Boards) eingesetzt. Auch diese Gremien setzen sich ausschließlich aus international renommierten Fachexperten zusammen und unterbreiten „auf Grundlage von externen Gutachten dem Kuratorium des FWF einen Entscheidungsvorschlag in Form eines Rankings“<sup>22</sup>. Auf Basis der Begutachtungsergebnisse trifft das Kuratorium des FWF die endgültige Entscheidung.

---

<sup>18</sup> Vgl. FWF, Evaluation im FWF, 1; <http://www.fwf.ac.at/de/downloads/pdf/evaluation-fwf.pdf> (Stand 26.7.2013).

<sup>19</sup> FWF, Förderungsrichtlinien, 1.

<sup>20</sup> Programm zur Entwicklung und Erschließung der Künste (PEEK); die Projektlaufzeit beträgt 36 Monate.

<sup>21</sup> Programm zur Förderung der Klinischen Forschung (KLIF); die Projektlaufzeit ist unterschiedlich.

<sup>22</sup> Vgl. FWF, Allgemeine Prinzipien des Entscheidungsverfahrens, 2013, 7; [www.http://www.fwf.ac.at/de/projects/entscheidungsverfahren/fwf-entscheidungsverfahren.pdf](http://www.fwf.ac.at/de/projects/entscheidungsverfahren/fwf-entscheidungsverfahren.pdf) (Stand 25.7.2013).



### *Interim-Evaluierungen während der Projektlaufzeit*

Der FWF führt bei Großprojekten in den Schwerpunkt- und Doktoratsprogrammen Evaluierungen zu den Projektfortschritten durch *Review-Panels* durch. So wird z.B. bei Doktoratskollegs durch Zwischenevaluierungen die Qualität des Projektfortschritts „sowohl im Hinblick auf den ‚Grundstock‘ wie auch im Hinblick auf erfolgte Verbreiterungen und Vernetzungen“<sup>23</sup> überprüft und am Ende jeder Förderperiode<sup>24</sup> über eine Verlängerung eines Fortsetzungsantrages entschieden.<sup>25</sup>

### *Ex post-Evaluierung (Qualitätssicherung)*

Die im Jahre 2003 eingeführte *ex post*-Evaluierung soll Informationen zum Projektverlauf sowie zum Projekterfolg liefern. Dabei werden Indikatoren wie „wissenschaftliche Veröffentlichungen und Beiträge bei Fachtagungen, Karriereverläufe von Projektbeteiligten (insbesondere des wissenschaftlichen Nachwuchses), besondere Auszeichnungen und Preise im Zusammenhang mit dem Projekt sowie anwendungsorientierte Resultate wie Patentanmeldungen, Lizenzen und Maßnahmen zur Wissenschaftskommunikation“<sup>26</sup> in der Projektdatenbank des FWF erfasst; diese Daten werden bei weiteren Förderentscheidungen berücksichtigt.

Zusätzlich zur Begutachtung von Förderanträgen und Projekten ist der FWF auch Auftraggeber von Evaluierungen; dies betrifft die regelmäßige Evaluierung der eigenen Förderprogramme durch internationale, unabhängige, einschlägig ausgewiesene Institutionen und die Evaluierung seiner Arbeitsweise als Organisation. Die Ergebnisse dieser ‚Metaevaluierungen‘ sind für generelle Aussagen zur Positionierung der Grundlagenforschung in Österreich relevant; sie erlauben auch Schlussfolgerungen zu den Effekten des nationalen Forschungsförderungs- und Wissenschaftssystems.

---

<sup>23</sup> [http://www.fwf.ac.at/de/applications/w/DK\\_Kurzinformation.pdf](http://www.fwf.ac.at/de/applications/w/DK_Kurzinformation.pdf), 2 (Stand 31.7.2013).

<sup>24</sup> Die Förderperiode für Doktoratskollegs beträgt jeweils vier Jahre.

<sup>25</sup> [http://www.fwf.ac.at/de/applications/w/DK\\_Kurzinformation.pdf](http://www.fwf.ac.at/de/applications/w/DK_Kurzinformation.pdf), 2 (Stand 31.7.2013).

<sup>26</sup> FWF, Evaluation im FWF, 1; <http://www.fwf.ac.at/de/downloads/pdf/evaluation-fwf.pdf> (Stand 26.7.2013).

### 2.2.2 Beispiel *Institute for Science and Technology* (IST Austria)

Das IST Austria ist ein österreichisches Forschungsinstitut mit Promotionsrecht<sup>27</sup>; es dient der Grundlagenforschung in den Bereichen Naturwissenschaften, Mathematik und Computerwissenschaften. Wissenschaftler sind nicht in Fachabteilungen, sondern in voneinander unabhängigen, interdisziplinären Forschungsgruppen organisiert, die von einem Professor oder einem Assistant Professor geleitet werden.

#### *Evaluierung des IST Austria*

Die Qualitätssicherungsmaßnahmen des IST Austria sind gesetzlich verankert; alle vier Jahre soll das Institut evaluiert werden. Die Evaluierungsberichte sind dem Nationalrat im Wege der Bundesregierung vorzulegen.<sup>28</sup> Im Jahre 2011 wurde die erste Evaluierung für den Finanzierungszeitraum 2007-2010 durchgeführt. Das Gutachterkomitee wurde, beruhend auf Vorschlägen des institutseigenen Wissenschaftlichen Rates zusammengestellt, unter dem Gesichtspunkt, dass es über „erhebliche Erfahrungen im Wissenschaftsmanagement“<sup>29</sup> verfügt. Laut Evaluierungsbericht lag das Hauptaugenmerk vorerst weniger auf der Begutachtung der Forschungsleistung des Instituts als vielmehr auf der Analyse von Strukturen, Regeln und Prozessen sowie der Qualität des wissenschaftlichen Personals.<sup>30</sup> Die im Bericht empfohlene langfristige Finanzierung wurde 2012 von Bund und Land Niederösterreich bis 2025 zugesagt.<sup>31</sup> Kernelemente einer Leistungsvereinbarung zwischen IST Austria und *bmwf/bmfwf* liegen bereits vor; die Vereinbarung soll bis spätestens 2016<sup>32</sup> unterzeichnet werden.

---

<sup>27</sup> Bundesgesetz über das Institute of Science and Technology Austria (2006) § 2 Abs. 1.

<sup>28</sup> [http://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA\\_2006\\_I\\_69/BGBLA\\_2006\\_I\\_69.pdf](http://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2006_I_69/BGBLA_2006_I_69.pdf), § 5 (Stand 10.9.2013).

<sup>29</sup> IST Austria, Evaluation of the Institute of Science and Technology Austria 2011, 3.

<sup>30</sup> A.a.O., 17. Die Evaluierung wurde in der Öffentlichkeit kritisch diskutiert; den Kernpunkt der Kritik bildete die Praxis der eigenen Auswahl des Gutachterkomitees.

<sup>31</sup> Das IST Austria erhält in einem Zeitraum von 2017 bis 2025 vom Bund jährlich im Schnitt 65,8 Mio. Euro sowie maximal 16,6 Mio. Euro leistungsabhängiger Förderung; das Land Niederösterreich stellt ein Budget von insgesamt 368 Mio. Euro für Infrastruktur, Gebäude und Campusbetrieb zur Verfügung (vgl. Pressemitteilung des IST Austria vom 22.2.2012).

<sup>32</sup> Parlamentarische Anfrage Nr. 11206/J-NR/2012 betreffend „Ungleichbehandlung der Finanzierung der Grundlagenforschung in Österreich am Beispiel des IST Austria“ vom 25. Mai 2012, 2.

### *Evaluierung der einzelnen Research Areas*

In den Jahren zwischen einer Gesamtevaluierung des IST Austria wird jährlich jeweils ein Fachbereich (*research area*) mittels *peer review*-Verfahren evaluiert. Die Entscheidung, welcher Bereich vertiefend evaluiert werden soll, liegt beim Wissenschaftlichen Rat des IST Austria, der auch die fachspezifischen, internationalen Gutachter auswählt. Das IST Austria übermittelt den Gutachtern – wie im Falle der Evaluierung des Gesamtinstitutes – einen Bericht zur Vorbereitung des Evaluierungsverfahrens, der organisatorische Informationen, strategische Ziele und Perspektiven für Fachbereiche sowie die wissenschaftlichen Lebensläufe und Forschungsleistungen der Mitarbeiter betrifft. Als für die Beurteilung des *research output* relevant werden Publikationen, Konferenzbeiträge, Patente und Softwareprodukte bezeichnet, in einem weiteren Sinne auch die Ausbildung und Karrierewege junger Forscher. Auch Ehrungen, Preise und Grants gelten als Indikator für wissenschaftlichen Prestigeerfolg und damit für die Qualität von Forschungsleistungen. Bibliometrische Indikatoren werden lediglich ergänzend herangezogen. Für die Beurteilung eines Fachbereichs werden ferner Professoren, Studenten, Postdocs und wissenschaftliche Mitarbeiter durch die Gutachter zu ihrer wissenschaftlichen Arbeit befragt; insofern ähnelt die Vorgangsweise der des britischen *Research Excellence Framework*. Der Abschlussbericht der Gutachtergruppe (des Panels) soll einen Vergleich zu den besten Forschungsgruppen erlauben und wird dem Wissenschaftlichen Rat sowie dem Präsidenten des IST Austria vorgelegt. Das Ergebnis beeinflusst interne Fördervergaben und Karrierepfade (Tenure).

### *Evaluierung von Assistant Professors*

Für die Funktion des Gruppenleiters rekrutiert das IST Austria Professoren im Rahmen einer unbefristeten Anstellung und Assistant Professors im Rahmen einer sechsjährigen Befristung. Neben jährlichen Überprüfungsgesprächen mit dem Präsidenten des IST Austria wird die wissenschaftliche Leistung des Assistant Professor nach fünf Jahren durch mindestens sechs internationale Gutachten (*review letters*) evaluiert. Die Gutachter werden vom *Professorial Committee* ausgewählt<sup>33</sup>; welches für Berufungsverfahren und Tenure Evaluierungen zuständig ist. Ob dem Assistant

---

<sup>33</sup> Das IST Austria hat Richtlinien zur gutachterlichen Beurteilung entwickelt, um (nach eigenen Angaben) Interessenskonflikte zu vermeiden. Vgl. D. Klammer, *Assessment and Measure of the Research Output at the Institute of Science and Technology Austria*, Klosterneuburg 2013, 1.

Professor nach positiver Evaluierung eine unbefristete Stelle als Professor angeboten wird, hängt von seiner in der Zeit am IST Austria erlangten Reputation in der internationalen *scientific community*, seinen Leistungen in Lehre und Mentoring und der Höhe der akquirierten Drittmittel ab.

### 2.2.3 Beispiel *Österreichische Akademie der Wissenschaften (ÖAW)*

Die Österreichische Akademie der Wissenschaften ist mit rund 1.300 Mitarbeiterinnen und Mitarbeitern und 28 Forschungseinrichtungen (FEs) „die führende außeruniversitäre Trägerinstitution für unabhängige, anwendungsoffene Grundlagenforschung in Österreich“. Zugleich bildet ihre Gelehrtengesellschaft mit ihren Mitgliedern aus dem In- und Ausland ein „nahezu alle Disziplinen umspannendes Diskursforum“<sup>34</sup>. Laut der erstmals mit dem *bmwf/bmfwf* abgeschlossenen Leistungsvereinbarung (LV) 2012-2014 steht der ÖAW für diesen Zeitraum ein Globalbudget von rund 223 Mio. Euro zur Verfügung.<sup>35</sup> In der LV werden wissenschaftliche Schwerpunkte gesetzt; das Forschungsportefeuille besteht nicht nur aus traditionellen Schwerpunkten, die ÖAW orientiert sich vielmehr auch an den Zielen der FTI-Strategie.

Die Messung und Beurteilung von Forschungsleistungen findet auf mehreren Ebenen statt: (1) Die FEs der ÖAW werden in regelmäßigen Abständen durch ein dafür geschaffenes Komitee evaluiert, (2) die Leistungen einzelner Mitarbeiter werden zusätzlich durch eine eigene Arbeitsgruppe beurteilt und (3) die Institute der ÖAW werden durch *Scientific Advisory Boards* (SABs) beraten und durch ein Monitoring begleitet.

#### *Evaluierung der Forschungseinrichtungen*

Unter der Prämisse, „die Wissenschaft auf allen Gebieten, insbesondere im Bereich der Grundlagenforschung, in jeder Hinsicht zu fördern“<sup>36</sup>, legt die ÖAW regelmäßig Evaluierungsberichte<sup>37</sup> vor. Diese dienen zum einen der Rechtfertigung für die Förderung aus öffentlichen Mitteln, zum anderen stellen sie eine Rückmeldung an die FEs über die Qualität ihrer Arbeit dar. Mit dem ersten Evaluationsturnus, dem „Mittel-

---

<sup>34</sup> Auszug aus dem Mission Statement der ÖAW, vgl. [www.oeaw.ac.at](http://www.oeaw.ac.at) (Stand 10.12.2013).

<sup>35</sup> Vgl. [http://www.bmwf.gv.at/uploads/tx\\_contentbox/Leistungsvereinbarung\\_2012-2014.pdf](http://www.bmwf.gv.at/uploads/tx_contentbox/Leistungsvereinbarung_2012-2014.pdf) (Stand 10.12.2013).

<sup>36</sup> MFP 2001-2005, 9f.

<sup>37</sup> <http://www.oeaw.ac.at/deutsch/forschung/qualitaet.html> (Stand 27.12.2013).

fristigen Forschungsprogramm 1996-2000“ (MFP)<sup>38</sup>, begann ein fortlaufender Prozess der Forschungsbewertung und Forschungsplanung.

Der Evaluierungsprozess folgt einem bestimmten Ablauf: Alle vier bis sechs Jahre ernennt das Präsidium der ÖAW, unterstützt durch in- und ausländische Fachkollegen, einen wissenschaftlich anerkannten Evaluierungsleiter, der in alleiniger Verantwortung ein akademieunabhängiges internationales Expertenkomitee für einen bestimmten Fachbereich<sup>39</sup> zusammenstellt und leitet. Die Evaluierung befasst sich jeweils mit einem Forschungsbereich und dem Beitrag einzelner Institutionen zum Gesamtbereich. Um die erforderliche Objektivität im Blick auf die zu evaluierende Forschungseinrichtung zu gewährleisten, besteht das Komitee ausschließlich aus externen Experten. Es prüft die dem Fachbereich angehörigen FEs und deren Arbeitsprogramme hinsichtlich ihrer mittelfristigen Forschungsplanung und zukünftigen Forschungsoptionen. Die MFPs bilden einen „Mechanismus der permanenten Qualitätssicherung“, der die Grundlage für Fortsetzung, Neuorientierung oder Beendigung der betreffenden Forschungsarbeiten“ darstellt<sup>40</sup>. So wurde nach Abschluss des MFP 1996-2000 z.B. die Forschungsstelle für Sozioökonomie geschlossen und das Institut für Informationsverarbeitung aufgelöst.

Die Einsetzung der Evaluationskomitees ist dem Forschungskuratorium<sup>41</sup> der ÖAW überantwortet. Seine Aufgabe ist es, die Leistung einer FE, eines Instituts<sup>42</sup>, im internationalen Vergleich zu begutachten. Die ausschließlich externe Zusammensetzung des Komitees wird „für jede Evaluation als empfehlenswert, auch für andere Instituti-

---

<sup>38</sup> Bislang wurden die Institute in drei Zyklen MFP 1996-2000, MFP 2001-2005 und MFP 2008-2012 evaluiert, im Zeitraum 2009-2012 zudem Zentren und 2010-2013 Cluster der ÖAW. Vgl. <http://www.oeaw.ac.at/deutsch/forschung/qualitaet.html> (Stand 29.11.2013).

<sup>39</sup> Ein Komitee evaluiert mit seiner fachlichen Expertise FEs aus dem gleichen Fachbereich. Die Heterogenität des Komitees erlaubt eine Beurteilung auf größerer disziplinärer Breite. Das MFP gibt schließlich Auskunft über die Forschungsplanung in den Fachbereichen der derzeit an der ÖAW existierenden Einrichtungen.

<sup>40</sup> MFP 2001-2005, 7.

<sup>41</sup> In strategischen Entwicklungsfragen und zur Sicherung der Fortentwicklung und Qualität der an der ÖAW betriebenen Forschung steht der ÖAW ein internationales Forschungskuratorium zur Seite. Vgl. <http://www.oeaw.ac.at/deutsch/about/forschungskuratorium.html> (Stand 29.11.2013).

<sup>42</sup> Vgl. § 70 (1) der Geschäftsordnung: Fachlich einander nahestehende Forschungseinrichtungen einer Klasse oder der Gesamtakademie können zu einem wissenschaftlichen Zentrum zusammengefasst werden; (...) (3) Die Evaluation von Forschungseinrichtungen eines Zentrums erfolgt im Rahmen der Evaluation des Zentrums.

onen“<sup>43</sup> angesehen. Nach Beendigung der Evaluierung löst sich das Komitee wieder auf; es ersetzt nicht die bestehenden Kontrollorgane der ÖAW.

### *Leistungsbewertung einzelner Mitarbeiter und Fächer*

Die Bewertung der Leistungen einzelner Mitarbeiterinnen und Mitarbeiter unterscheidet sich von den anderen Evaluierungen der ÖAW. Für die Beurteilung von Anträgen auf Entfristung zeichnet eine eigene Arbeitsgruppe verantwortlich. Was die laufende Evidenz wissenschaftlicher Leistung betrifft, werden in den Naturwissenschaften die üblichen quantitativen Kriterien wie etwa *h-Faktor* oder *Science Citation Index* mit herangezogen. Für die geisteswissenschaftlichen Fächer wurden an der ÖAW bereits 2008 erste Überlegungen angestellt, auf die gestützt konkrete Kriterien erarbeitet werden sollen.<sup>44</sup>

### *Scientific Advisory Boards (SABs)*

Jedes einzelne ÖAW-Institut wird von einem wissenschaftlichen Beirat, dem *Scientific Advisory Board* (SAB) begleitet. Seine Aufgabe besteht darin, das Institut und die ÖAW-Leitung in Fragen der bestmöglichen strategisch-wissenschaftlichen Weiterentwicklung zu beraten. Dazu zählt auch das Monitoring der Arbeit der einzelnen Gruppen und deren Leiter. Der SAB tritt im Laufe seiner Amtsdauer nach Bedarf zusammen. Im Sinne der akademischen Aufgabentrennung<sup>45</sup> werden die SABs von der Gelehrten-gesellschaft eingesetzt. Der SAB kann bis zu 50 Prozent aus Österreichern bestehen, bei geisteswissenschaftlichen Instituten mit starkem Österreichbezug kann dieser Anteil auch höher sein.<sup>46</sup>

Der SAB des Instituts für Schallforschung z.B. beschreibt in einem „Report“ die Struktur und Mission des Instituts; er evaluiert in diesem Report sowohl das Institut in seiner Gesamtheit als auch die einzelnen Forschungsgruppen – wobei die Publikationsleistung, bewilligte Drittmittel und Projektgelder (z.B. des FWF), internationale Koope-

<sup>43</sup> A. Zeilinger, Präsident der ÖAW in einem E-Mail an den ÖWR am 16.10.2013.

<sup>44</sup> Vgl. <http://www.oeaw.ac.at/gema/PDF/BewertungGeisteswissAdhoc08.pdf> (Stand 10.12.2013).

<sup>45</sup> Vgl. ÖAW Jahresbericht 2012: „Im Oktober 2012 fasste die Gesamtsitzung der ÖAW den Grundsatzbeschluss zur Entflechtung der Gelehrten-gesellschaft und der Forschungsträgerinstitution unter dem gemeinsamen Dach der Akademie. (...) Die Gelehrten-gesellschaft bleibt dem Forschungsträger in diversen Funktionen wissenschaftlich beratend verbunden, sie wird jedoch von operativen Aufgaben seines Managements befreit, um sich auf die Pflege des interdisziplinären wissenschaftlichen Diskurses konzentrieren und dessen Ergebnisse künftig verstärkt in die unabhängige, wissenschaftsbasierte Gesellschafts- und Politikberatung einbringen zu können.“

<sup>46</sup> Vgl. A. Zeilinger in seiner E-Mail an den ÖWR am 16.10.2013.

rationen, die aktive Mitwirkung an der Forschungsentwicklung auf dem jeweiligen Gebiet etc. die Bewertung beeinflussen. Abschließend reiht der SAB die einzelnen Gruppen in deskriptiver Form, Beispiel: „The SAB ranks this group as between excellent and outstanding“. Am Ende gibt der Beirat noch eine Empfehlung ab: „the ARI should continue in its present form (...)“<sup>47</sup>

Der SAB des Instituts für Molekulare Biotechnologie (IMBA) z.B. setzt sich ausschließlich aus international renommierten Wissenschaftlern zusammen und trifft sich jährlich, um die Qualität und die internationale Signifikanz der Forschung zu bewerten und Vorschläge für die weitere Ausrichtung des Instituts zu diskutieren.<sup>48</sup>

Die Arbeit des SAB ist für jedes Institut besonders vor dem internen Hintergrund wichtig. Bei der Programmevaluation durch das einmalig eingesetzte Evaluierungskomitee wird hingegen betont, dass hier die Transparenz zwischen den ÖAW-Instituten gefördert und die wissenschaftlichen Leistungen offengelegt werden. Zudem soll die Evaluierung auch dazu dienen, das Zukunftspotential der FEs, Zentren, Cluster oder Kommissionen vor dem internationalen Hintergrund auszuloten und strategische Schritte zu setzen.

---

<sup>47</sup> <http://www.kfs.oeaw.ac.at/pub/ari11final.pdf> (Stand 02.12.2013).

<sup>48</sup> Vgl. <http://de.imba.oeaw.ac.at/ueber-imba/struktur-organisation/management/wissenschaftlicher-beirat/> (Stand 02.12.2013).

### 3. Stellungnahme und Empfehlungen

#### 3.1 Allgemeine Einschätzungen

Messung und Beurteilung wissenschaftlicher Leistungen sind in individueller wie institutioneller Hinsicht mit vielen Schwierigkeiten behaftet. Grund ist vor allem das ungeheure Wachstum des Systems Wissenschaft, das hier nach neuen Instrumenten, neben dem älteren Instrument des *peer review*, ruft; es sind aber auch die wachsenden Anforderungen an die Qualitätssicherung in der Wissenschaft.

Zu den neuen Instrumenten gehören, wie dargestellt, quantitative Methoden, vor allem in bibliometrischer Form. Diese Methoden sind mittlerweile in vielen, vor allem in den empirischen Wissenschaften, allen voran in den Naturwissenschaften, fest etabliert, stoßen aber, abgesehen von einer generellen kritischen Einschätzung, die sich auf eine zu starke Betonung quantitativer Gesichtspunkte bezieht, in den Geisteswissenschaften, in den nicht-empirischen Formen der Sozialwissenschaften und in der Forschung in den Künsten auf erhebliche Probleme. Diese bestehen darin, dass die Arbeitsweisen andere, oft sehr individuell geprägte sind, dass (in der Terminologie Th. Kuhns) unterschiedliche Paradigmen die wissenschaftliche Arbeit bestimmen und es meist keine qualitativ abgestuften Publikationsformen gibt. Diese allgemeinen wie disziplinspezifischen Probleme sind bekannt und finden in der Regel im Beurteilungszusammenhang Berücksichtigung.

Unter den Untersuchungen, die sich mit Empfehlungen zur Messung und Beurteilung von Forschungsleistungen befassen, sind auch solche des deutschen Wissenschaftsrates. Demnach können „Veränderungen in der Forschungspraxis, die häufig auf Verfahren der Steuerung und Bewertung von Forschungsleistung zurückgeführt werden, vielfältige und vor allem auch systembedingte Ursachen haben (...), wie etwa: Das Größenwachstum des Wissenschaftssystems, die wachsenden externen Ansprüche an die Wissenschaft, die Internationalität der Forschung, (...) die zunehmende Angewiesenheit auf Drittmittel, die nicht immer gegebene Nachhaltigkeit wissenschaftspolitischer Initiativen, die Strukturierung des Systems in einen hochschulischen und einen außeruniversitären Sektor.“<sup>49</sup> Diese Veränderungen haben „dazu beigetragen, Leistungsunterschiede in der Forschung transparenter und nachvoll-

---

<sup>49</sup> Empfehlungen zur Bewertung und Steuerung von Forschungsleistung, Köln 2011, 5.



ziehbarer zu machen<sup>50</sup>. Maßgebend ist auch für den deutschen Wissenschaftsrat das Verhältnis von quantitativen und qualitativen Methoden bei der Messung und Beurteilung von Forschungsleistung. So werden die Vorteile der Verwendung quantitativer Indikatoren darin gesehen, dass diese „in erster Linie der Komplexitätsreduktion und der Ermöglichung von Vergleichbarkeit (dient). Quantitative Indikatoren erfüllen den Bedarf nach Messbarkeit und lassen sich vergleichsweise gut operationalisieren. Zudem kann durch eine formelbasierte Verrechnung quantitativer Indikatoren der Aufwand für Gutachterinnen und Gutachter reduziert werden.“<sup>51</sup> Nachteile werden wieder darin gesehen, dass die Verwendung von Indikatoren ‚manipulationsanfällig‘ ist und zu ‚Zielverschiebungen‘ führen kann, womit „schließlich nicht mehr die Verbesserung der Forschungsleistung das Ziel ist, sondern die quantitative Steigerung im Hinblick auf bestimmte Indikatoren.“<sup>52</sup> Hinsichtlich der Verwendung qualitativer Methoden wird festgestellt, dass die Leistungsbewertung „durch *peers* in sogenannten *peer review*-Verfahren (...) im Vergleich der Bewertungsverfahren am ehesten ein fundiertes Qualitätsurteil über Forschung und eine institutionalisierte wissenschaftsinterne (Selbst-)Kontrolle (erlaubt). Sie wird der Vielfalt und Komplexität der Forschung am ehesten gerecht.“<sup>53</sup> Andererseits wird konstatiert, dass das Begutachtungsverfahren in einem „System permanenter und eng getakteter Bewertungsprozesse (...) an seine Grenzen (stößt), da der Aufwand für alle Beteiligten hoch ist.“<sup>54</sup>

Eine Entscheidung zugunsten der einen Methode (indikatorengesteuertes Verfahren) oder der anderen (*peer review*-Verfahren) wird in der Untersuchung des deutschen Wissenschaftsrates nicht getroffen, argumentiert wird für eine Verbindung beider Methoden. So wird empfohlen, sich über fachinterne Qualitätsstandards zu verständigen<sup>55</sup>, werden quantitative Methoden hinsichtlich eines ergänzenden Charakters betont und wird eine Interpretation von Indikatoren durch qualifizierte *peers* gefordert, „um fachbezogene Einordnung unter Berücksichtigung der spezifischen Funktion und methodischen Konstruktion des Indikators zu gewährleisten“<sup>56</sup>. „Da Indikatoren For-

---

<sup>50</sup> A.a.O., 15.

<sup>51</sup> A.a.O., 19.

<sup>52</sup> Ebd.

<sup>53</sup> A.a.O., 16.

<sup>54</sup> Ebd.

<sup>55</sup> A.a.O., 38.

<sup>56</sup> A.a.O., 42.

schungsleistung lediglich anzeigen, ohne sie zu erklären, ist stets eine Interpretation durch qualifizierte Personen erforderlich.<sup>57</sup> Aus einem Gegensatz wird auf diese Weise ein Ineinandergreifen zweier Methoden, wobei disziplinspezifische Unterschiede eine wesentliche Rolle spielen.

In seinen neuesten „Empfehlungen zur Zukunft des Forschungsratings“ (2013) setzt sich der deutsche Wissenschaftsrat dann ausdrücklich für ein „qualitatives *peer review*-Verfahren“<sup>58</sup> ein, das fächerbezogen und über unterschiedliche Institutionalisierungen hinweg (Universitäten, außeruniversitäre Forschungseinrichtungen, Fachhochschulen) Forschungsleistungen beurteilbar machen soll. Dazu wird ein Verfahren vorgeschlagen, das ausgehend von einer fächerspezifischen Operationalisierung (Bestimmung des Fachgebietes und seiner Teile, Entwicklung eines spezifischen Bewertungsmodells) über Datenerhebung und Datenanalyse zur Bewertung führt.<sup>59</sup> Dieses Verfahren soll selbst in Form flächendeckender Vergleiche in allen zu ‚Bewertungsblöcken‘ zusammengefassten Disziplinen institutionalisiert werden, was wiederum einen großen institutionellen und organisatorischen Aufwand bedeutet. Hier droht die Gefahr, dass sich ein Verfahren verselbständigt, dass es eher sich selbst genügt, als die Forschung in Qualität und Leistungsfähigkeit zu fördern.

Das sieht auch der Schweizerische Wissenschafts- und Innovationsrat in einer Untersuchung des Verhältnisses von indikatorengestützter (quantitativer) und gutachtergestützter (qualitativer) Leistungsbeurteilung so.<sup>60</sup> Dabei wird kritisch darauf hingewiesen, dass ein Übermaß an quantitativen, insbesondere bibliometrischen, Verfahren das gesuchte Leistungsprofil über die dabei Anwendung findenden Reduktionen (gegebener Mehrdimensionalität) verzerren kann, wie auch ein Übermaß an drittmittelorientierter Bewertung störende Nebeneffekte wie Mainstreamphänomene zur Folge haben kann. Umgekehrt wird auf Schwächen eines *peer review*-Systems hingewiesen, die unter anderem in der häufigen Widersprüchlichkeit von Gutachteraussagen und einer Routinisierung von gutachtengestützter Evaluierung lägen. Bezogen auf die Verbindung von *informed peer review*-Verfahren und Evaluierungsdichte wird festgestellt: „Die Lösung dieses Dilemmas, selbst in der *Informed Peer Review* keine wirklich zufriedenstellende, nachhaltige Form der Leistungsoffenlegung

---

<sup>57</sup> A.a.O., 39.

<sup>58</sup> Köln 2013, 37.

<sup>59</sup> A.a.O., 39ff..

<sup>60</sup> Leistungsmessung und Qualitätssicherung in der Wissenschaft, SWTR-Schrift 3/2013, Bern 2013.

zu finden, liegt (...) in der Anpassung der Frequenz, Form und Zielsetzung der Evaluationsverfahren. Selbst wenn die Leistungsbeurteilung einer hohen Fehlermarge unterworfen ist, können Evaluationen dennoch zur Qualitätsentwicklung beitragen, allerdings nur unter bestimmten kommunikativen Bedingungen, in denen ein offener Dialog und die Offenlegung der eigenen Schwächen als Teil eines Versuches, diese zu beheben und die Stärken besser nutzen zu können, gelingen. Gerade bei evaluationsgeprüften Forschungseinheiten bringen weniger routinisierte, d.h. seltenere Evaluationen mit einem Fokus auf Fragen, die sich die Einheit tatsächlich auch selbst stellt, mehr Hoffnung auf einen anregenden, wirklich auf Qualitätsverbesserung zielenden Evaluationsprozess.<sup>61</sup> Betont wird in diesem Zusammenhang die positive Rolle einer *ex ante*-Qualitätssicherung (Ausbildung und Berufung) gegenüber der üblichen Form einer *ex post*-Kontrolle.

Im übrigen bezieht sich das hier zum Stichwort ‚Evaluation‘ (bzw. ‚Evaluierung‘) Gesagte auch auf andere institutionelle Formen der Leistungsbeurteilung und Qualitätssicherung im wissenschaftlichen Bereich, so in modifizierter Form auf *Audits* und *Akkreditierungen*. Evaluierungen werden dabei allgemein als „systematische Anwendung sozialwissenschaftlicher Forschungsmethoden zur Beurteilung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens sozialer Interventionsprogramme“ definiert.<sup>62</sup> Sie werden weit gefasst und betreffen die Untersuchung sowohl von Prozessen als auch von Resultaten. Sie können auf individueller, programmatischer oder institutioneller Ebene durchgeführt werden. Klassische Evaluationsverfahren im Hochschulbereich umfassen drei Phasen: eine Phase der kritischen Selbstreflexion (interne Evaluation), eine Phase des *peer review* (externe Evaluation) und eine Phase der Rückkoppelung der Resultate (Umsetzung). Alle Phasen werden durch entsprechende Berichte dokumentiert. Zur Illustration dieser Verfahren und ihrer Unterschiede die folgende Tabelle:

---

<sup>61</sup> A.a.O., 38.

<sup>62</sup> P. H. Rossi/H. E. Freeman/G. Hofmann, Programm-Evaluation. Einführung in die Methoden angewandter Sozialforschung, Stuttgart 1988, 3.

Verfahren	dient in erster Linie der	Untersuchungsgegenstand	Instrumente	Ebene	Mindeststandards
Evaluation	Qualitätssicherung	Prozesse, Ergebnisse	1, 2	individuell, programmatisch, institutionell	nein
Akkreditierung	Qualitätssicherung	Prozesse, Ergebnisse	1, 2	programmatisch, institutionell	ja
Audits	Qualitätssicherung	Prozesse	1, 2	institutionell	nein
Leistungsvereinbarungen	Steuerung und Rechenschaftslegung	Zielvereinbarungen	2	individuell, institutionell	

Tab 1: 1 = Quantitative Methoden: Bibliometrie, andere Leistungsindikatoren, die nicht standardisierbar sind, weil dazu keine referentiellen sets existieren, Wissensbilanz, Rankings.  
2 = Qualitative Methoden: *peer review*.

Quantitative und qualitative Instrumente zur Messung und Beurteilung von Forschungsleistungen, sofern sich diese im wesentlichen auf publizierte Leistungen beziehen, stehen, wie in Kapitel 1 und auch hier unter dem Stichwort ‚Evaluation‘ näher dargestellt, nicht allein. Sie werden in der neueren Evaluationspraxis von weiteren Maßnahmen begleitet. Diese betreffen den zuvor wie auch in der Stellungnahme des Schweizerischen Wissenschafts- und Innovationsrates erwähnten Drittmittelsektor, ferner den Promotions- und Habilitationssektor und ein so genanntes *informed peer review*.<sup>63</sup> So werden Drittmittelinwerbungen z.B. im Förderranking der Deutschen Forschungsgemeinschaft (DFG) und in den Hochschulrankings des Centrums für Hochschulentwicklung (CHE) als Dokumentation von Forschungsleistung gesehen, und zwar sowohl hinsichtlich ihrer Anzahl als auch hinsichtlich ihrer Volumina. Gemessen wird in diesem Falle nicht eine Outputgröße, wie im Falle eines bibliometrischen Vorgehens, sondern eine Inputgröße. Von ihr wird auf eine besondere For-

<sup>63</sup> Dazu auch: T. Grapat u.a., Kriterien zur Messung der Forschungsleistung an der Heinrich-Heine-Universität Düsseldorf, 36 ff..

schungsintensität geschlossen. Mit den Drittmitteln wird ein Produktionsfaktor gemessen, nicht ein Erkenntnisfortschritt, der der eigentliche Maßstab einer Beurteilung von Forschungsleistungen sein sollte. Dabei könnte ein hohes Maß an Abhängigkeit von Drittmitteln geradezu kontraproduktiv in Hinblick auf das eigentliche Erkenntnisziel sein, nämlich dann, wenn der Forscher zu großen Teilen seiner Arbeit zum Drittmittelmanager wird.

Das gleiche gilt für den Bereich der Promotionen und Habilitationen. Diese bilden in ihrer gängigen Beurteilung die Intensität forschungsorientierter Betreuungsleistungen ab, nicht die Qualität von Promotionen und Habilitationen. Im Gegensatz dazu werden in *informed peer review*-Verfahren ausgewählte Publikationen (einzelner Wissenschaftler oder ganzer Forschungseinrichtungen) hinsichtlich ihrer Qualität und Originalität von *peers* beurteilt, womit qualitative Maße und Gesichtspunkte in den Vordergrund rücken. Diese erfassen neben quantitativ erhobenen Daten auch Stärken und Schwächen individueller wie institutioneller Art, die sich einer rein quantitativen Betrachtung entziehen, ohne damit gleich vor subjektiven Fehleinschätzungen und widersprüchlichen Beurteilungen, wie auch in Kapitel 1 beschrieben, gefeit zu sein. Der Grund besteht hier darin, dass es keine international anerkannten Standards gibt, nach denen sich *peers* verlässlich richten können. Diese sind teilweise auf ihre persönlichen Erfahrungen angewiesen und bringen sie bewusst oder unbewusst in eine Evaluierung ein. Daher ist bei der Auswahl der *peers* auch eine besondere Sorgfalt geboten, und dies insbesondere dann, wenn es sich nicht nur um fachspezifische, sondern auch um institutionelle Fragestellungen handelt. In einem solchen Zusammenhang sind Kenntnisse der nationalen Rahmenbedingungen erforderlich. Wird ihnen zu wenig Aufmerksamkeit gewidmet – und das gilt auch für Beurteilungen individueller Qualität –, verlieren entsprechende Verfahren an Validität und Reliabilität.

Auch die genannten ergänzenden Maßnahmen und Instrumente führen also gewisse Nachteile mit sich. Drittmittelmaße benachteiligen Forschungsbereiche, die stärker durch individuelle Leistungen geprägt und nicht geräteintensiv sind, Promotions- und Habilitationsmaße lassen in ihrer quantitativen Verwendung keine verlässlichen Rückschlüsse auf Qualität zu, *informed peer review*-Verfahren, obgleich am ehesten geeignet, die Anforderungen an eine Leistungsmessung und eine Qualitätsbeurteilung zu erfüllen, sind aufwendig und anfällig für subjektive Einschätzungen.

Fazit ist auch hier, dass der Königsweg zur Messung und Beurteilung von Qualität in der Forschung in der Verbindung sowohl qualitativer Verfahren (*peer review*) als auch quantitativer Verfahren (z.B. bibliometrischer Verfahren) liegt. Beide Verfahren sollten sich wechselseitig ergänzen, wobei das jeweilige Maß vom Gegenstand der Beurteilung abhängt: nicht alle disziplinären Bereiche sind gleicher Art (Beispiel: die Natur- und die Geisteswissenschaften) und nicht alle institutionellen Bereiche folgen den gleichen Aufgaben (Beispiel: Universitäten und Fachhochschulen). Und selbst innerhalb gleicher disziplinärer Bereiche sind fächerspezifische Besonderheiten gegeben, die sich auch in der Auswahl der Verfahren niederschlagen sollten (Beispiel: die Archäologie und die Literaturwissenschaft in den Geisteswissenschaften). Den Ausgangspunkt aller Messungs- und Beurteilungsversuche sollte insofern die Beantwortung folgender Fragen bilden: Was wird in einem nach innen wie nach außen differenzierten Wissenschaftssystem gemessen und beurteilt: Dasselbe? Mit denselben Instrumenten? Muss nicht je nach Aufgabe und Selbstverständnis gemessen und beurteilt werden? Nach Gegenstand und nach institutioneller Aufgabe?

Sind diese Fragen geklärt, könnte im Sinne eines *informed peer review*-Verfahrens die Regel (jedenfalls im Rahmen eines Evaluationsverfahrens) sein: ein qualitatives Vorgehen, gestützt durch quantitative Indikatoren, wo immer sich dies nahelegt. Auch hier geht es also um eine kombinierte Vorgangsweise, in der alle zuvor genannten Faktoren wie Publikationen, Drittmittel, Promotions- und Habilitationsintensitäten, aber auch Auszeichnungen, Kooperationen und herausgehobene Funktionen in der *scientific community* berücksichtigt werden, alles in allem aber hinter einer primär qualitätsorientierten und eher holistischen Beurteilung zurückstehen sollten. Dies mag im Einzelfall schwierig sein und oft auch kontrovers ausfallen, ist aber immer noch die beste Lösung, wenn es wirklich darum geht, Qualität, die im Forschungszusammenhang nie einfach zu beschreiben ist, zu beurteilen. Würde man sich – was leider nur zu oft der Fall ist – auf die Feststellung von *impact*-Faktoren beschränken, wäre es um die wissenschaftliche Beurteilungsfähigkeit, damit auch um die wissenschaftliche Objektivität im Beurteilungszusammenhang geschehen.

### 3.2 Österreichspezifische Empfehlungen

Die Situation in Österreich unterscheidet sich nach Vorgehen und herrschender Beurteilungskultur, wie in Kapitel 2 dargestellt, nicht wesentlich von anderen Wissenschaftssystemen. Besonderheiten sind z.B. das hohe Maß an Autonomie, das den Universitätssektor auszeichnet, und das Instrument der Leistungsvereinbarung, das in besonderem Maße geeignet ist, unterschiedliche Bemessungs- und Beurteilungsformen, darunter auch wieder qualitative wie quantitative Verfahren, zusammenzuführen. Das gilt in vergleichbarer Weise auch für die außeruniversitären Forschungseinrichtungen und für den forschenden Fachhochschulbereich. In all diesen Zusammenhängen ist auch die Möglichkeit gegeben, in einem Beurteilungszusammenhang Rücksicht auf fachliche und disziplinäre Besonderheiten zu nehmen. Zu diesen gehören z.B. die besondere Stellung von Monographien im Leistungsausweis der Geisteswissenschaften (wobei allerdings auch hier, wie zuvor erwähnt, ein allmählicher Prozess hin zu einer stärkeren Gewichtung von Einzelbeiträgen und damit hin zu *peer review*-Verfahren zu beobachten ist) und die Rezensionskomponente/Kritikkomponente in den Kunstwissenschaften, deren Praxis einer forschungsbezogenen Leistungsbeurteilung, ein weiter Forschungsbegriff unterstellt, im übrigen der der Geisteswissenschaften in mancher Hinsicht ähnelt.

Insofern läuft auch eine Empfehlung zur Förderung und Sicherung der Forschungsqualität sowie zum Umgang mit den entsprechenden Instrumenten einer Leistungsmessung und Leistungsbeurteilung für Österreich auf ein *informed peer review*-Verfahren unter Berücksichtigung fächer- und disziplinspezifischer Unterschiede hinaus. Es sollte nach Ziel und Vorgehen das Standardverfahren darstellen, und zwar auf zwei Ebenen: (1) der Ebene der Bewertung der Forschungsleistung österreichischer Forschungseinrichtungen im Vergleich, vor allem als Grundlage für die Leistungsvereinbarungen, (2) der Ebene interner Evaluierungen als Grundlage strategischer Entscheidungen der Forschungseinrichtungen, z.B. im Kontext einer gesuchten Profil- und Schwerpunktsetzung und bei der inneruniversitären Mittelallokation. Beide Ebenen hängen aufs engste miteinander zusammen, und zwar im Idealfall so, dass jede Universität – für außeruniversitäre Forschungseinrichtungen in ähnlicher Weise – ein valides, an den hier beschriebenen Standards orientiertes Verfahren der Bewertung durchführt und die Ergebnisse dieser Form der Evaluierung einen Vergleich zwischen den Universitäten erlauben. Erforderlich dazu wäre, dass

eine Standardisierung des hier vorgeschlagenen Verfahrens eines *informed peer review* unter Berücksichtigung fächer- und disziplinspezifischer Unterschiede sich, wie auch vom deutschen Wissenschaftsrat empfohlen, auf Fächer- und Disziplinen- gruppen bezieht und für diese ein Anforderungsprofil zu einem solchen Verfahren oder ein entsprechendes ‚Drehbuch‘ existiert, auf das sich die Akteure zuvor ver- ständigigt haben.

Der Wissenschaftsrat empfiehlt die Entwicklung derartiger ‚Drehbücher‘ im Zu- sammenspiel von Universitäten und zuständigem Ministerium – ein entsprechender Prozess im außeruniversitären Forschungsbereich müsste wohl gesondert angesto- ßen werden, weil hier die Ausgangsbedingungen andere sind –, wobei es ratsam wäre, hier zunächst mit einem Pilotprojekt, etwa für eine naturwissenschaftliche und eine geisteswissenschaftliche Fächer- oder Disziplinengruppe, zu beginnen.

Empfohlen sei in diesem Zusammenhang nur noch, das Instrument der Wissensbi- lanz zurückzufahren. Dieses stellt ein ausschließlich quantitatives Verfahren zur Lei- stungsmessung dar und bietet nur wenig Handhabe für eine wirkliche Qualitätsbeur- teilung. Worum es auch hier geht, ist die Förderung einer strategischen Schwer- punkt- und Profildisziplinbildung, mit der ein qualitatives Umfeld geschaffen wird, dessen auch die wirklich herausragende (‚exzellente‘) Forschung bedarf. Ein *informed peer review*-Verfahren unter Berücksichtigung fächer- und disziplinspezifischer Unter- schiede in den Dienst einer solchen Förderung zu stellen, würde den gesamten Mess- und Beurteilungsverfahren einen substantiellen Sinn geben und nicht nur der Legitimationsbeschaffung dienen, wie häufig anzutreffen. Das war zuvor mit der Ge- fahr einer Verselbständigung der Instrumente zur Leistungsmessung und Lei- stungsbeurteilung gemeint, die einem Wissenschaftssystem nicht dient, sondern es zum Objekt ihrer Selbstdarstellung nimmt.

Kurzgefasst wird für die österreichische Situation, unter Einbeziehung einiger weite- rer Gesichtspunkte, empfohlen:

### 3.2.1 Im Allgemeinen:

- Entscheidung für eine enge Verbindung sowohl quantitativer als auch qualitativer Verfahren im Sinne eines *informed peer review*-Verfahrens unter Berücksichti- gung fächer- und disziplinspezifischer Unterschiede. Eine Berücksichtigung



dieser Unterschiede ist geboten, weil es kein Mess- und Beurteilungsverfahren gibt, das in gleicher Weise in allen Fächern und Disziplinen Anwendung finden könnte. Dazu sind die Wissenschaftskulturen zu verschieden – nicht nur aus historischen, sondern auch aus systematischen Gründen. Beispiele: Naturwissenschaften – Geistes- und Kulturwissenschaften, empirische – nicht-empirische Wissenschaften.

- Beachtung des Unterschieds zwischen individueller und institutioneller Leistungsmessung und Leistungsbeurteilung (Qualitätsbeurteilung). Das Leistungsprofil eines einzelnen Forschers ist unter Heranziehung sehr unterschiedlicher, zum Teil wiederum individueller wie institutioneller Art anders zu beurteilen als das einer Institution.
- Zeiträume zwischen der Durchführung von Leistungsmessungen und Leistungsbeurteilungen (Leistungsvereinbarungen, Evaluierungen, Audits, Akkreditierungen etc.) weiter fassen, damit Messung und Beurteilung bzw. die Vorbereitung für derartige Prozesse der Wissenschaft und ihrer Qualität wirklich dienen und nicht zu deren Lasten, in den Universitäten zu Lasten von Forschung und Lehre, gehen. Empfohlen wird, wie auch von den Universitäten so gesehen, eine Fünfjahresspanne anstelle der häufig, z.B. in den Leistungsvereinbarungen, praktizierten Dreijahresspanne. Es muss vermieden werden, dass mit Abschluss eines Prüfungsprozesses schon wieder die Vorbereitungen für den nächsten Prozess beginnen (müssen).
- Stärkere Berücksichtigung von Konsequenzen durchgeführter Leistungsmessungen und Leistungsbeurteilungen sowohl aus der Betroffenen- als auch aus der Prüfendenperspektive. Prüfungsprozesse gehen ins Leere bzw. bleiben ohne sichtbare Spuren, wenn sie sich gegenüber der Praxis, z.B. der Hochschulpraxis, der sie dienen sollen, verselbständigen, d.h. zum Selbstzweck werden. Als positives Beispiel für folgenreiche Evaluierungen kann die Ludwig Boltzmann Gesellschaft (LBG) mit der Schließung von Instituten nach erfolgter negativer Evaluierung dienen.
- Entlastung des Gutachtersystems. Ein großer Teil des Forschungspotentials ist heute im Begutachtungswesen gebunden. Das fängt bei der Zahl der Gutachter in den Promotionen an, setzt sich bei der Zahl der Gutachter bei beantragten For-

schungsprojekten fort und endet in großen Gutachtergremien auf regionaler und nationaler Ebene. Das Gutachtersystem stößt schon lange an seine Grenzen und geht mittlerweile zu Lasten des Forschungssystems („schleichende Systemüberlastung“, FWF). Dabei wird oft im Quantitativen schon das Qualitative gesehen, dient der hohe Aufwand an Begutachtungsressourcen eher Legitimations- als Qualitätsgründen. Mehr Aufmerksamkeit sollte statt einer möglichst großen Zahl von Gutachtern der Qualität einer kleinen Zahl von Gutachtern gewidmet werden.

- Neujustierung des Drittmittelwesens. Drittmittel, zumal wenn sie der Förderung der Spitzenforschung dienen sollen und hochkompetitiv vergeben werden, stellen ein wichtiges Element der Forschungsförderung dar. Sie verlieren diese Bedeutung, wenn sie zu einem großen Teil der Kompensation sonst fehlender Mittel dienen und ihre reine Quantität schon als Ausweis wissenschaftlicher Leistungsfähigkeit angesehen wird. Hier gilt es sich auf die ursprüngliche Funktion derartiger Mittel zu besinnen und das Wissenschaftssystem entsprechend neu zu justieren. Für Österreich heißt das, die chronische Unterfinanzierung der Universitäten aufzuheben, um zu verhindern, dass Drittmittel zweckentfremdet für die Grundbedürfnisse von Forschung und Lehre eingesetzt werden, zugleich aber die Mittel des Fonds zur Förderung der wissenschaftlichen Forschung (FWF) für die Spitzenforschung wesentlich zu erhöhen. Nur so kann die Einwerbung von Mitteln in Beurteilungszusammenhängen wieder als Indikator für wirkliche Spitzenforschung dienen.

### 3.2.2 Im Besonderen:

- Akkreditierungshürden höher legen. Angesichts der österreichischen Hochschulwirklichkeit stellt sich die Frage, ob die Akkreditierungshürden im Wissenschaftsbereich nicht zu niedrig liegen und damit zu einer Nivellierung des wissenschaftlichen und institutionellen Niveaus führen. Diese ist z.B. dort gegeben, wo privaten Einrichtungen, die zum Teil nicht einmal ein Fach oder eine Disziplin in allen ihren Teilen abdecken oder nur in geringem Maße über ein überzeugendes Forschungspotential verfügen, der Status einer Universität zukommt. Gegebenenfalls sind hier entsprechende gesetzliche Maßnahmen zu treffen.

- Wissenschaftliche Grade und entsprechende Anforderungen an die Ausbildung schärfen. Wie im Falle der herrschenden Akkreditierungspraxis muss auch bei den akademischen Abschlüssen gewährleistet sein, dass sie internationalen Standards entsprechen. Das gilt insbesondere im österreichischen Promotionswesen. Hier sind Tendenzen zu einem *Dr. light* erkennbar, d.h. einer Promotion, die kein hohes Forschungsniveau und kein breites fachliches und disziplinäres Umfeld, wie international üblich, zur Voraussetzung hat (Promotionsrecht der privaten Universitäten, neuerdings auch der Donau-Universität Krems, Diskussion über ein Promotionsrecht der Fachhochschulen).
- Leistungsvereinbarung und Leistungsbeurteilung als allgemeines Modell. Das Instrument der Leistungsvereinbarung hat sich im Universitätsbereich im wesentlichen bewährt. Es sollte, auch über die Erweiterung im Falle der ÖAW und des IST Austria hinaus, in allen forschenden (und lehrenden) Einrichtungen, die durch öffentliche Mittel finanziert (oder mitfinanziert) werden, Anwendung finden, also auch im Bereich der außeruniversitären Forschung generell, der privaten Universitäten (sofern diese auch öffentlich gefördert werden) und der Fachhochschulen (als lehrende Einrichtungen mit begrenztem Forschungsauftrag). Dies würde auch der Einheitlichkeit eines in sich differenzierten Wissenschaftssystems dienen.

## **ÖSTERREICHISCHER WISSENSCHAFTSRAT**

**IMPRESSUM** Medieninhaber und Herausgeber: Republik Österreich/Österreichischer Wissenschaftsrat, Liechtensteinstraße 22a, 1090 Wien, Tel.: 01/319 49 99-0, Fax: 01/319 49 99-44, [office@wissenschaftsrat.ac.at](mailto:office@wissenschaftsrat.ac.at), [www.wissenschaftsrat.ac.at](http://www.wissenschaftsrat.ac.at) Umschlaggestaltung: Starmühler Agentur & Verlag, [www.starmuehler.at](http://www.starmuehler.at) Druck: Gerin