PLATTFORM ▪ TECHNOLOGIE ▪ EVALUIERUNG

Nr. 10 Juni 2000

■ THEMA

Evaluierung von Forschungsinstitutionen in Deutschland
Hans-Dieter Daniel

Evaluation of University Research: Strategies for Austria
David F. J. Campbell

Evaluation of the Austrian Academy of Sciences
Armin Scrinzi

Evaluation of Biomedical Research in Austria
Bernd R. Binder

■ BERICHTE

Evaluierungsseminar Karlsruhe
Tagungsbericht

Eine Initiative von
BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE
JOANNEUM RESEARCH FORSCHUNGSGES.M.B.H.
ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG
ÖSTERREICHISCHES FORSCHUNGSZENTRUM SEIBERSDORF
TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H.

bmvit JOANNEUM RESEARCH WIFO AUSTRIAN RESEARCH CENTERS TiG Technologie Impulse Gesellschaft m.b.H.

Vorwort

Nach einer längeren Pause halten Sie nun wieder ein – sehr umfangreiches - Exemplar des Newsletters der PLATTFORM TECHNOLOGIE EVALUIERUNG in Händen. In Abstimmung mit der im Jänner dieses Jahres durchgeführten Plattform-Diskussionsveranstaltung ist dieses Heft der Evaluierung im Forschungsbereich gewidmet. Das Thema wird dabei von zwei unterschiedlichen Blickwinkeln aus betrachtet: Zum einen aus der Sicht eines Evaluierungsexperten und eines Forschungsökonomens; zum anderen aus der Sicht von Wissenschaftlern, die aktiv in die Evaluierung „ihrer" Forschungsinstitutionen involviert waren.

Im ersten Teil des Newsletters bietet Hans-Dieter Daniel einen Überblick über die Evaluierungsaktivitäten in Deutschland und David Campbell analysiert mögliche Strategien für die Implementierung von Evaluierungen an österreichischen Universitäten. Danach kommen die „Praktiker" zu Wort: Armin Scrinzi schildert seine Erfahrungen mit der Evaluierung der Österreichischen Akademie der Wissenschaften, während Bernd Binder die Entwicklung einer Evaluierungs-„Best Practice" im Bereich der österreichischen Biomedizinforschung beschreibt.

Ein Kurzbericht von Michael Stampfer über die Ergebnisse eines kürzlich in Karlsruhe abgehaltenen Evaluierungsseminars beschließt diesen Newsletter.

Viel Spaß beim Lesen!

Preface

After a longer than expected while we can now offer you a new edition of the Newsletter of the PLATTFORM TECHNOLOGIE EVALUIERUNG, which is devoted to evaluations in the field of science and research. This topic is viewed from two different angles: First the issue is approached by two experts in science and research evaluation and policy; afterwards, two scientists who were actively involved in the evaluation of their own research institutions describe their experience.

In the first part of the Newsletter Hans-Dieter Daniel provides an overview and assessment of past and current research evaluation activities in Germany, while David Campbell analyses potential strategies for the implementation of evaluations standards at Austrian universities. Then Armin Scrinzi provides his experience from the evaluation of the Austrian Academy of Science and Bernd Binder describes the development of Best Practice evaluation standards in the field of biomedicine in Austria.

Finally Michael Stampfer summarises the results of an evaluation seminar held recently in Karlsruhe.

Enjoy reading this Newsletter, yours

Oliver Fritz

on behalf of the Plattform-Team

Evaluierung von Forschungs-institutionen in Deutschland

Hans-Dieter Daniel

Modernisierungsdruck und Sparzwang haben dazu geführt, dass immer mehr Länder Verfahren für die Evaluierung von Forschungseinrichtungen entwickeln und einsetzen. Seitdem der quantitative Ausbau des Forschungssystems in vielen Ländern zum Stillstand gekommen ist (vgl. Goodstein, 1996) und sich in der Situation des "Steady State" (Ziman, 1987) neue Forschungseinrichtungen nur durch Beendigung alter oder veralteter Arbeitsfelder und durch Schließung von Abteilungen oder Instituten realisieren lassen, ist die Suche nach "objektiven Bewertungsmaßstäben" weltweit intensiviert worden.

Eine Recherche aktueller Veröffentlichungen zum Thema macht deutlich, dass die Evaluierung von Forschung heute ein ubiquitäres Phänomen ist. Das von Daniel & Fisch (1986) für den Zeitraum von 1910 bis 1985 konstatierte exponentielle Wachstum der Literatur zum Themenkomplex Forschungsevaluation hat sich auch in den 90er Jahren weiter fortgesetzt. Während zu diesem Themenkomplex zwischen 1910 und 1949 durchschnittlich weniger als eine Arbeit pro Jahr publiziert wurde, waren es zwischen 1950 und 1959 im Durchschnitt drei Arbeiten pro Jahr, zwischen 1960 und 1969 zehn Arbeiten pro Jahr, zwischen 1970 und 1979 29 Arbeiten pro Jahr und im Zeitraum von 1980 bis 1985 durchschnittlich 35 Arbeiten pro Jahr. Für die letzten 15 Jahre liegen keine genauen statistischen Angaben vor. Eine von Kostoff (1998a) zusammengestellte Bibliographie neuerer Arbeiten zum Themenbereich Forschungs- und Technologieevaluation umfasst 5600 Titel, eine weitere zum "Peer Review" allein 1500 Veröffentlichungen (Kostoff, 1998b).

In Deutschland werden etwa seit 1975 Ansätze zur Messung und Beurteilung von Forschungsleistungen entwickelt und erprobt (für eine Übersicht vergleiche Fisch & Daniel, 1986; Daniel & Fisch, 1988; Daniel, 1989 sowie Daniel & Fisch, 1990). Eine der ersten Evaluationsstudien wurde von Klingemann (1974) durchgeführt, der für Wissenschaftler der Kernforschungsanlage Jülich den Zusammenhang zwischen Publikations- und Zitationshäufigkeiten analysierte. Im Auftrag des damaligen Bundesministeriums für Bildung und Wissenschaft verglich Spiegel-Rösing (1975) fächerübergreifend die Forschungsleistungen bundesdeutscher Universitäten. Als Indikatoren für Forschungsleistung verwandte sie ebenfalls Publikations- und Zitationsmaße. Spiegel-Rösing (1975) erstellte erstmals auch Universitäts-Ranglisten nach US-amerikanischem Vorbild. Ranglisten für die deutschen Universitäten hatte es bis dahin nicht gegeben, weil sich die Hochschulpolitik der Nachkriegszeit am Grundsatz der Gleichbehandlung orientiert hatte. Gemeinhin war man davon ausgegangen, dass alle Universitäten gleichrangig seien. Das änderte sich Anfang der 70er Jahre, als zunehmend Kritik am Leistungsniveau der deutschen Forschung geäußert (vgl. Maier-Leibnitz, 1979 sowie May, 1997 und 1998 für einen aktuellen Effizienzvergleich nationaler Forschungssysteme und Herbertz & Müller-Hill, 1995 für einen internationalen Effizienzvergleich molekularbiologischer Forschungsinstitute) und die Finanzierung der öffentlichen Haushalte schwieriger wurde. Differenzierung, Profilbildung, Transparenz, Evaluation und Wettbewerb sind seither Schlüsselbegriffe der forschungspolitischen Diskussion.

In den 80er Jahren entwickelte sich in Deutschland eine empirische Wissenschaftsforschung, die unterstützt vom Stifterverband für die Deutsche Wissenschaft, der Deutschen Forschungs-gemeinschaft und den damaligen Bundesministerien für Bildung und Forschung, zahlreiche weitere Evaluationsstudien

durchgeführt hat, und zwar sowohl fächerübergreifend als auch auf einzelne Fachgebiete bezogen (vgl. Daniel, 1989 und Daniel & Fisch, 1990). In diesen überwiegend auf ein Ranking von Universitäten abzielenden Studien wurden sowohl bibliometrische Indikatoren als auch eine Reihe von Reputationsmaßen erprobt. Neben der Befragung von Wissenschaftlern (Peers) und Führungskräften aus Wirtschaft, Politik und Verwaltung waren dies vor allem ortsvergleichende Auswertungen der Förderstatistiken der Alexander von Humboldt-Stiftung und der Deutschen Forschungsgemeinschaft (die DFG hat 1997 erstmals selbst eine Drittmittelstatistik, aufgeschlüsselt nach Hochschulen und vier Fächergruppen, vorgelegt. Eine Fortschreibung der Statistik ist in Vorbereitung).

Bei den Evaluationsstudien der 80er Jahre handelte es sich überwiegend um bibliometrische Untersuchungen, d. h. es wurden die Produkte der Forschung mit Hilfe statistischer Verfahren analysiert, wie beispielsweise die Verteilung von Veröffentlichungen, Patenten und Zitationen auf Wissenschaftler, Forschungseinrichtungen und Länder (für neuere deutsche Studien in dieser Tradition vgl. Creutzfeldt & Gerok, 1997; Frömmel & Heß, 1998; Vitt, 1998). Nur vereinzelt wurde versucht, den wissenschaftlichen Input in Form von Personal, Geräten und Verbrauchsmaterial zu erfassen und ihn zum Forschungsoutput in Beziehung zu setzen (für eine der wenigen Effizienzanalysen vgl. Herbertz & Müller-Hill, 1995). Erste Erfahrungen wurden auch mit sogenannten Mapping-Techniken gesammelt, die mit Hilfe der Ko-Zitations-Cluster-Analyse versuchen zu ermitteln, welche Forschungsgebiete – unabhängig von möglicherweise veralteten Disziplineinteilungen – zur Zeit am intensivsten erforscht werden und welche Wissenschaftler, Forschungseinrichtungen und Nationen auf diesen Gebieten führend tätig sind (vgl. Daniel, 1988 und Small, 1999 für neuere Entwicklungen im Bereich der Ko-Zitations-Cluster-Analyse). Darüber hinaus gab es erste Versuche, Indikatoren für die Internationalität der Forschung – ein Thema, das in der aktuellen wissenschaftspolitischen Diskussion in Deutschland einen hohen Stellenwert hat - zu entwickeln (vgl. Münzinger & Daniel, 1992b und 1994).

Die Ergebnisse dieser Evaluationsstudien, die sehr stark mit Fragen der Methodenentwicklung befasst waren (Leistungsindikatoren, Datenbanken, statistische Analyseverfahren), lassen sich wie folgt zusammenfassen:

Fächerübergreifende Leistungsvergleiche sind wegen der gravierenden disziplinären Unterschiede (wie etwa Drittmittelaufkommen, Publikations- und Zitationsgepflogenheiten) äußerst schwierig. Vor diesem Hintergrund verdient die fach- und institutionenübergreifende Analyse und Bewertung großer Forschungsfelder durch den Wissenschaftsrat besondere Erwähnung: Mit seiner im Jahr 1994 verabschiedeten Stellungnahme zur Umweltforschung in Deutschland hatte der Wissenschaftsrat zum ersten Mal eine Querschnittsstudie zu einem großen Forschungsfeld vorgelegt, das über die Grenzen der klassischen Fachdisziplinen und der Sektoren der öffentlich finanzierten Forschungseinrichtungen hinweg bewertet wurde (vgl. Wissenschaftsrat, 1994). Für die sich anschließende Querschnittsbegutachtung der Materialforschung wurde ein zweistufiges Verfahren gewählt. Zunächst wurden “Empfehlungen zur Förderung materialwissenschaftlicher Forschung und Lehre an den Universitäten” vorgelegt, die sich vor allem mit den Strukturen material-wissenschaftlicher Forschung und Lehre an den Universitäten in Deutschland befassen. Darauf aufbauend wurde die “Stellungnahme zur außeruniversitären Materialwissenschaft” ausge-arbeitet, in der mehr als 30 Einrichtungen der Materialforschung dargestellt und bewertet

wurden (vgl. Wissenschaftsrat, 1999a). Mit der Stellungnahme zur Energieforschung liegt die dritte und vorläufig letzte Querschnittsbewertung vor, wobei universitäre und außeruniversitäre Einrichtungen gemeinsam betrachtet werden (vgl. Wissenschaftsrat, 1999b).

Bei der Bildung bibliometrischer Indikatoren ist darauf zu achten, dass Spezifika der Datenbanken (vgl. Münzinger & Daniel, 1992a; Quoniam et al., 1995) und der auf Indikatorenbildung spezialisierten Institute (z. B. ISI, CHI, ISRU) die Ergebnisse nicht verzerren bzw. verfälschen (vgl. Glänzel, 1996).

Evaluationen sollten sich auf einen Zeitraum von drei bis fünf Jahren beziehen, weil kürzere Zeiträume instabile, längere Zeiträume veraltete Daten liefern.

Die Evaluierung von Leistungen in der Forschung sollte anhand eines Indikatorenbündels erfolgen. Die Evaluationskriterien sollten sich auf Umfang und Qualität der Forschung (vgl. Buchholz, 1995) beziehen und neben subjektiven Bewertungen immer auch (quasi-)objektive Daten berücksichtigen. Unverzichtbar für die Evaluation von Forschungsleistungen ist das Urteil der Fachkollegen (Peer Review). Reliabilität (im Sinne von Gutachterübereinstimmung), Fairness und prädiktive Validität des Peer-Review-Verfahrens sind allerdings nicht voll befriedigend (vgl. Neidhardt, 1988; Daniel, 1993). Um die Urteilsgrundlage der Gutachter zu verbessern, sollten ihnen zumindest Informationen über (1.) Anzahl und Wirkung von Zeitschriftenpublikationen, (2.) Buchpublikationen und Buchbeiträge, (3.) Kongress- und Tagungsaktivitäten, (4.) Höhe der eingeworbenen Drittmittel und (5.) Ämter und Berufungen zur Verfügung gestellt werden. Diese quasi-objektiven Evaluationskriterien sind, wie eine Befragung unter Psychologie-Professoren in Deutschland gezeigt hat, zumindest innerhalb der Profession konsensfähig (vgl. Montada, Krampen & Burkard, 1999).

Evaluationsergebnisse sollten in Form eines Leistungsprofils dargestellt werden, weil nur diese Präsentationsform Stärken und Schwächen der Forschungseinrichtung erkennen lässt. Diese Informationen können von der Forschungsinstitution im Sinne einer formativen Evaluation genutzt werden, indem sie beispielsweise dazu anregen, über geeignete Organisations- und Personalentwicklungsmaßnahmen nachzudenken.

Ein fairer Leistungsvergleich von Forschungseinrichtungen sollte immer auch bestehende Größen- und Ausstattungsunterschiede berücksichtigen, damit nicht fälschlicherweise als Leistung ausgewiesen wird, was in Wirklichkeit nur Ausdruck der Größe oder der personellen und infrastrukturellen Ausstattung der Institutionen ist (vgl. den in dieser Hinsicht vorbildlichen Leitfaden für die Bewertung von Einrichtungen der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibnitz des Wissenschaftsrates, 1999c).

Da eine absolute Quantifizierung von Forschungsleistungen nicht möglich ist, kann die Höhe der Leistung immer nur im Vergleich zu anderen “Leistungserbringern” bestimmt werden. Der Auswahl der Vergleichsobjekte (z. B. nach dem Benchmarking-Ansatz) kommt daher eine besondere Bedeutung zu.

Da die Leistungsunterschiede innerhalb der Forschungsinstitutionen häufig wesentlich größer sind als die Leistungsunterschiede zwischen den Forschungseinrichtungen, sollte im Rahmen von Evaluationen immer auch ermittelt werden, welche Wissenschaftler die eigentlichen Leistungsträger sind (vgl. Narin & Hamilton, 1996; Vitt, 1998). Dieser Punkt ist ganz besonders wichtig, weil keine Forschungseinrichtung als Ganzes einem Kriterium für Spitzenforschung genügen kann. Forschungsergebnisse werden immer von Personen erarbeitet, die sich hinsichtlich ihrer Produktivität und Reputation in hohem Maße unterscheiden. Innerhalb jeder

Einrichtung existieren ganz erhebliche Leistungsunterschiede zwischen den einzelnen Wissenschaftlern (in der Regel in der Größenordnung von eins zu zehn). “Das Leben der Universität hängt”, um Karl Jaspers (1961) zu zitieren, “an den Persönlichkeiten, nicht an der Institution, welche nur Bedingung ist” (S. 124). Dies gilt nicht nur für Universitäten, sondern für alle Forschungs-einrichtungen gleichermaßen.

Während in den 70er Jahren in Deutschland die Einschätzung überwog, dass Forschungsleistungen nicht evaluierbar seien, hat in den letzten Jahren ein Meinungswandel stattgefunden. Dies mag sicherlich auch mit der normativen Kraft des Faktischen zusammenhängen: evaluationsfreie Zonen sind im deutschen Forschungssystem kaum noch existent. Angesichts der zunehmenden Zahl von Evaluationen sprechen einige Beobachter bereits von “Evaluitis” und betrachten die sich abzeichnende Evaluationsbürokratie im “evaluative state” (vgl. Neave, 1998) bzw. die Legitimationsrituale der “audit society” (vgl. Power, 1997) mit großer Skepsis.

Abstract

The literature on research evaluation is growing exponentially. The main reasons for this development are the end of expansion in science budgets, the increasing value of basic research for commercial technology, and the intensified technological competition in global markets. This paper reviews the history of research performance evaluation in Germany over a 25-year period from 1975-99.

Literatur

Buchholz, K. (1995). Criteria for the analysis of scientific quality. Scientometrics, 32(2), 195-218.

Creutzfeldt, W. & Gerok, W. (Hrsg.). (1997). Medizinische Publizistik: Probleme und Zukunft. Stuttgart: Georg Thieme Verlag.

Daniel, H.-D. (1988). Evaluation der biotechnologischen und molekulargenetischen Forschung in der Bundesrepublik Deutschland anhand des ISI Atlas of Science - Eine Momentaufnahme bedeutender Forschungsbeiträge. In: H.-D. Daniel & R. Fisch (Hrsg.), Evaluation von Forschung (S. 121-150). Konstanz: Universitätsverlag Konstanz.

Daniel, H.-D. (1989). Ansätze zur Messung und Beurteilung des Leistungsstandes von Forschung und Technologie: Eine Bestandsaufnahme biblio-, sziento- und technometrischer Forschungen in der Bundesrepublik Deutschland. Teil 1: Stand und Perspektiven der Forschung. Teil 2: Bibliographie bundesdeutscher Beiträge. Beiträge zur Hochschulforschung, Heft 3, 223-253.

Daniel, H.-D. (1993). Guardians of Science - Fairness and Reliability of Peer Review. Weinheim: Wiley-VCH.

Daniel, H.-D. & Fisch, R. (1986). Messung von Forschungsleistungen - Eine annotierte Bibliographie (1910 - 1985) und Synopsis. Erlangen: Institut für Gesellschaft und Wissenschaft.

Daniel, H.-D. & Fisch, R. (Hrsg.). (1988). Evaluation von Forschung: Methoden - Ergebnisse - Stellungnahmen. Konstanz: Universitätsverlag Konstanz.

Daniel, H.-D. & Fisch, R. (Guest Editors). (1990). Scientometrics Research in the Federal Republic of Germany. Scientometrics, 19 (5-6), 327- 512.

Daniel, H.-D. & Fisch, R. (1990). Research Performance Evaluation in the German University Sector. Scientometrics, 19 (5-6), 349-361.

Deutsche Forschungsgemeinschaft (1997). Bewilligungen nach Hochschulen. Bonn.

Fisch, R. & Daniel, H.-D. (Hrsg.). (1986). Messung und Förderung von Forschungsleistung: Person - Team - Institution. Konstanz: Universitätsverlag Konstanz.

Frömmel, C. & Heß, D. (Hrsg.). (1998). Leistungsbewertung Forschung: Förderung der Forschung an Medizinischen Fakultäten und Hochschulen. Berlin: Klarsicht Verlag.

Glänzel, Wolfgang (1996). The need for standards in bibliometric research and technology. Scientometrics, 35(2), 167-176.

Goodstein, D. (1996). The big crunch: the end of expansion in science. Washington, D. C.: George C. Marshall Institute.

Herbertz, H. & Müller-Hill, B. (1995). Quality and efficiency of basic research in molecular biology: a bibliometric analysis of thirteen excellent research institutes. Research Policy, 24, 959-979.

Jaspers, K. (1961). Die Idee der Universität. In: K. Jaspers & K. Rossmann (Hrsg.), Die Idee der Universität (S. 41-165). Berlin.

Klingemann, H. (1974). Ein Beitrag zur Methode der Messung individueller wissenschaftlicher Leistung – dargestellt am Beispiel der Kernforschungsanlage Jülich. Zeitschrift für Soziologie, 3, 356-374.

Kostoff, R. N. (1998a). Science and Technology Metrics. Arlington, VA: Office of Naval Research.

Kostoff, R. N. (1998b). Research Program Peer Review: Principles, Practices, Protocols. Arlington, VA: Office of Naval Research.

Maier-Leibnitz, H. (1979). Doppelt soviel Spitzenforschung an den deutschen Hochschulen? In: H. Maier-Leibnitz, Zwischen Wissenschaft und Politik: Ausgewählte Reden und Aufsätze 1974-1979. Boppard: Harald Boldt.

May, R. M. (1997). The scientific wealth of nations. Science, 275 (7 February), 793-796.

May, R. M. (1998). The scientific investments of nations. Science, 281 (3 July), 49-51.

Montada, L., Krampen, G. & Burkard, P. (1999). Persönliche und soziale Orientierungslagen von Hochschullehrern/innen der Psychologie zu Evaluationskriterien über eigene berufliche Leistungen. Psychologische Rundschau, 50(2), 69-89.

Münzinger, F. & Daniel, H.-D. (1992a). Die Forschung in der ehemaligen DDR im Spiegel bibliometrischer Indikatoren: Möglichkeiten und Grenzen von Online-Datenbanken. In: W. Neubauer & K.-H. Meier (Hrsg.), Deutscher Dokumentartag 1991: Information und Dokumentation in den 90er Jahren - Neue Herausforderungen, neue Technologien (S. 303-319). Frankfurt am Main: Deutsche Gesellschaft für Dokumentation.

Münzinger, F. & Daniel, H.-D. (1992b). Biochemie und Molekularbiologie in der Bundesrepublik Deutschland: Bibliometrische Indikatoren 1973-1987. Futura - Informationen aus dem Boehringer Ingelheim Fonds, 7(4), 4-11.

Münzinger, F. & Daniel, H.-D. (1994). Biochemie und Molekularbiologie in der Bundesrepublik Deutschland: Internationale Ko-Autorenschaft und deren Einfluß auf die Zitierhäufigkeit von Veröffentlichungen 1973 bis 1987. Futura - Informationen aus dem Boehringer Ingelheim Fonds, 9(1), 20-25.

Narin, F. & Hamilton, K. S. (1996). Bibliometric performance measures. Scientometrics, 36(3), 293-310.

Neave, G. (1998). The evaluative state reconsidered. European Journal of Education, 33(3), 265-284.

Neidhardt, F. (1988). Selbststeuerung in der Forschungsförderung: Das Gutachterwesen der DFG. Opladen: Westdeutscher Verlag.

Power, M. (1997). The Audit Society: Rituals of Verification. Oxford: Oxford University Press.

Quoniam, L., Rostaing, H., Boutin, E. & Dou, H. (1995). Treating bibliometric indicators with caution: their dependence on the source database. Research Evaluation, 5(3), 177-181.

Small, Henry (1999). Visualizing Science by Citation Mapping. Journal of the American Society for Information Science, 50(9), 799-813.

Spiegel-Rösing, I. (1975). Zur Messung der Forschungsleistung von Institutionen: Wissenschaftliche Produktivität westdeutscher Universitäten. Eine explorative Untersuchung. In: Bundesministerium für Bildung und Wissenschaft (Hrsg.), Beiträge zur Messung von Forschungsleistung: Institutionen, Gruppen und Einzelpersonen (S. 15-80). Bonn: BMBW (Schriftenreihe Hochschule, Band 16).

Vitt, J. (1998). Schlüsselerfinder in der industriellen Forschung und Entwicklung. Wiesbaden: Deutscher Universitätsverlag.

Wissenschaftsrat (1994). Stellungnahme zur Umweltforschung in Deutschland (zwei Bände). Köln.

Wissenschaftsrat (1999a). Stellungnahme zur außeruniversitären Materialwissenschaft. Köln.

Wissenschaftsrat (1999b). Stellungnahme zur Energieforschung. Köln.

Wissenschaftsrat (1999c). Leitfaden für die Bewertung von Einrichtungen der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibnitz (WGL) (Blaue Liste). Drs. 3927/99. Köln.

Ziman, J. (1987). Science in a "Steady State" – The Research System in Transition. London: Science Policy Support Group (SPSG Concept Papers No. 1).

Autoren-Hinweis

Hans-Dieter Daniel, Wissenschaftliches Zentrum für Berufs- und Hochschulforschung, Universität Gesamthochschule Kassel, Henschelstr. 4, D-34109 Kassel,
Tel.: ++49-561-804-2419,
E-mail: daniel@hochschulforschung.uni-kassel.de

Evaluation of University Research: Strategies for Austria

David F. J. Campbell

The Importance of Research and University Research for Advanced Knowledge-Based Societies

For the contemporary knowledge-based societies and advanced economies, it is clear that processes of high-quality knowledge production should be regarded as crucially important. In such a conceptual understanding, high-quality knowledge production represents a key factor for determining the competitiveness — and consequently also the wealth — of the industrialised countries during their transition to "post-industrial" information societies (Bell, 1999). High-quality knowledge production is substantially realised through research or R&D (research and experimental development). Therefore, research must be regarded as pivotal for sustaining and improving the performance and competitiveness of advanced economies (Felderer & Campbell, 1994).

Within the context of national (and supranational) systems of innovation and research, the universities and university research express an important functionality. In most OECD countries, university research is primarily public funded and focuses on basic research. Thus universities also perform long-term research. In contrast to universities, business R&D is more application and experimental-oriented. Even though the total R&D expenditure of universities is considerably lower than that of business (the business enterprise sector), university research is not "less" important, since it "produces" basic research results. The shortening life cycles of economic products and services can be used as an argument, which again stresses the importance of basic research, since basic research potentially promises long-term commercial utilisation opportunities. And basic research mainly is located at the universities. For business companies this seems to imply two strategies: first, processes of "paralleling" of basic and applied research, experimental development, and market introduction; second, an emphasised linkage-building between business and universities.

The demand profile, put on university research, also changes. The concept of "Mode 2" attempts to summarise these trends. A team of researchers under the guidance of Michael Gibbons (*et al.,* 1994) invented and defined "Mode 2" as being based on the following five principles: "knowledge produced in the context of application"; "transdisciplinarity"; "heterogeneity and organisational diversity"; "social accountability and reflexivity"; and "quality control". For university research, particularly two implications appear to be important: first, the significance and "intrinsic scientific values" of basic research are increased, when basic research is performed within an application-oriented context. Second, the conceptualisation of quality of university research should be broadened by adding on additional dimensions to the "traditional" core concept of university quality; for example, efficiency, relevance, viability, and effectiveness (Campbell, 1999, 375-376; see also VSNU, 1994, 1998).

Evaluations as a Means of Optimising University Research

The importance of university research for the national systems of innovation and research creates a need for implementing structures, which attempt to permanently optimise university research. Such optimisation is necessary, independently of the question whether the resource input into university research increases or decreases. Evaluations are regarded as key tools or strategies that support this optimisation

by stressing the following functionality (Campbell, 1999, 370-372):

1. generating systematic feedback mechanisms;
2. supporting the implementation of market mechanisms into the university system; however, the rationale of those "academic markets" demands a specific adaptation to the needs of universities and may not be a simple copy of the rationale of "economic markets";
3. amplifying the rationality of the university system by defining "objective" criteria for resource allocation and individual career promotion: evaluation results qualify as criteria for decision-making;
4. justifying the use of public resources; universities and university research are primarily public funded, thus there is a need for universities to express accountability concerning the proper use of resources.

With regard to the methodology of university evaluations, two alternative approaches are possible: *indicators or (and/or) "peer review".* Indicators aim at data (information) that can be measured and collected quantitatively. Peer review, on the other hand, bases assessment on the judgement of "peers" (experts or scientists). Each of those approaches demonstrates specific strengths and weaknesses. The strengths are: for peer review "complexity" and for indicators "objectivity". Consequently, the weaknesses are: again, for peer review "subjectivity" and for indicators "superficiality". Elaborated in more detail, the following meaning is implied by these terminological concepts (Campbell, 1999, 374-375):

1. *Complexity:* Quality assessment based on expert judgement and expert panels covers, usually, a wider spectrum of information than a "pure" indicator system. Furthermore, experts are also in a position of conducting an analysis of advanced complexity.
2. *Subjectivity:* One of the crucial problems of peer review is the factor of subjectivity. Consequently, this refers to the following challenge: What are the control mechanisms, which prevent, that an assessment outcome is overly determined by a specific composition of panel members?
3. *Objectivity:* A main advantage of indicators should be seen in the circumstance that they are based on information, which can be measured, counted, and thus — in principle — is receptive for an "objective" ("inter-subjective") validation.
4. *Superficiality:* Are there any control mechanisms of indicator systems in place, which guarantee, that the information (or data), which is measured, also represents the relevant information? Furthermore, often a procedural policy problem exists of validating the collected quantitative data.

In practice, comprehensive evaluation systems regularly combine peer review and indicators. The evaluation of university research in the United Kingdom and the Netherlands can be described as being primarily organised as a peer review process, which is supported and underpinned by indicator-based information (Campbell, 1999, 376-378).

Strategies for the Evaluation of University Research in Austria

Strategies for evaluating university research in Austria but also in Germany (Campbell & Felderer, 1997) are challenged by the following key issues (Campbell, 1999, 378-381):

1. *Comprehensive institutional ex post evaluations:* Derived from the permanent need of optimising

university research, a comprehensive evaluation system on a nation-wide scale should be regarded as necessary for Austria. Taking into consideration that the financing of Austrian university research is primarily generated by public basic funds (so-called GUF), it follows that an institutional evaluation ought to apply the principles of an *ex post* evaluation (Felderer & Campbell, 1994). Currently, such a comprehensive and nation-wide evaluation system still neither exists nor operates in Austria (the same applies to Germany).

2. *Consequences of systematic ex post evaluations:* Certainly it should be considered a pivotal question, whether systematic linkages between evaluation outcome and the funding of institutions should be installed. One can argue convincingly that evaluation results ought to impose at least some consequences. In that respect, the spectrum of possible consequences extends broadly: ranging from symbolic incentives, over "real" incentives, to fundamental structural changes. Consequences concerning resource allocations may address the basic funding, the earmarked funding (for projects and programs) and university-internal re-allocations. Obviously, such issues always will be discussed, and determined, controversially.

3. *Combining the evaluation of research with the evaluation of teaching (study programs):* Universities are interpretable as institutions that fulfil a multi-functional profile. Clearly, research and teaching represent core responsibilities of universities. Conventionally, evaluations of research and teaching are separate procedures. This may be justified by referring to the resulting "over"-complexity when research and teaching should be evaluated simultaneously through one evaluation procedure. In principle, the methodical binary dichotomy of indicators and peer review also applies to the evaluation of teaching. When evaluations of research and teaching are carried out separately (at least partially), this produces the crucial challenge of re-integrating afterwards the results into a comprehensive conclusion that reflects the whole functional performance of a university. Systems for the evaluation of university research ought to have built-in interfaces for evaluations of university teaching.

4. *Combining evaluations of university and university-related research:* A growing functional overlapping between university and university-related (in German "außeruniversitäre") research can be stated. Universities and university-related institutions together define the so-called "academic research cluster". This creates a certain demand for a standardised "core" compatibility between research evaluation systems of university and university-related institutions, which would enable direct quality and efficiency comparisons. This puts, furthermore, an expectation on evaluation systems of university research to demonstrate, how they can be coupled with evaluation models for the university-related sectors.

Derived from such general considerations at the Institute for Advanced Studies (*Institut für Höhere Studien* / IHS) in Vienna, there is a model — which we will call in the following simply *"Evaluation Model"* — for the comprehensive evaluation of university research in Austria that was developed and proposed for discussion (Campbell & Felderer, 1999). This Model represents the final outcome of a research project that was commissioned by the Austrian Federal Ministry of Science and Transport (BMWV).

Although the Evaluation Model focuses primarily on university research, it also extends to university-related research, so that evaluation results for all academic institutions can be generated in a standardised and comparable format. However, the evaluation recommendations for the university-related research are "softer" and conceptually more explorative than the recommendations for universities. In addition, the proper status of this Evaluation Model should be regarded as that of a qualitative discussion contribution for the current evaluation discourse, as is being debated in Austria.

The university-oriented recommendations of the Evaluation Model apply the following principles:

1. *Combining indicators and peer review:* A dual evaluation mode is proposed that combines a monitoring system with systematic *ex post* evaluations. Both approaches are regarded as equal and, mutually linked, they should produce a synergy between the methodical advantages of "indicators" and "peer review" ("objectivity" and "complexity").

2. *Monitoring:* Based on primarily quantitative data and indicators (that include input and output), research should be regularly observed and "measured". Derived from such a monitoring, the research efficiency of universities and university departments may be calculated continuously.

3. *Ex post evaluation:* Through applying the method of peer review, the research quality of university departments (in accordance to their disciplinary assignment) can be evaluated by expert panels. Four different quality dimensions should be addressed: quality (in a "traditional core" understanding), efficiency, relevance, and viability. Furthermore, all departments ought to be graded on a five-point rating scale (1-5) for each dimension. The *ex post* evaluation of all university departments (disciplines) may be performed within a temporal evaluation cycle frame of 3-4 years.

4. *Research points:* "Research points" should be assigned to the university departments that reflect the outcome of monitoring and *ex post* evaluation. Through a bottom-up procedure those research points at the department level can be aggregated to research points for a whole university. This system of research points addresses two functions: first, displaying in public the combined evaluation results for the monitoring and *ex post* evaluations. Second, offering a "rational" basis for outcome-bound resource distributions. Furthermore, the Evaluation Model equally weights the research points for monitoring ("research efficiency") and *ex post* evaluations ("research quality"). Obviously, this symmetric emphasis may be changed by weighting monitoring or *ex post* evaluation differently. If, for instance, the peer review is considered as more important or valuable, then there ought to be a weighting in favour of *ex post* evaluation ("research quality").

Although the presented Evaluation Model focuses specifically on the Austrian university system (see again Campbell & Felderer, 1999), a discussion could be added, to which extent this Model also demonstrates applicability to university systems in other European countries. In Germany, for instance, the prime public responsibility for universities is not concentrated at the federal level (*Bund*), but is located sub-nationally at the province ("Federal State") level; the so-called *Länder* (Campbell & Felderer, 1997). This makes policy comparisons between individual

German *Länder* and the whole Austrian nation — concerning the "argument of size" — easier; also in the field of university research evaluation. Therefore, it would be interesting to test, whether the recommendations of the Evaluation Model could also be of some relevance for universities in Germany (or other Central-European countries).

Abstract

University research should be regarded as crucially important for the advanced national systems of innovation and research, since it focuses on basic research with a long-term perspective. Evaluations attempt optimising university research through the following mechanisms: generating feedback; defining academic markets; amplifying rationality; justifying public resource use. Methodically, two different evaluation approaches for university research can be distinguished: *indicators* and *peer review*. Both approaches express specific advantages ("objectivity" and "complexity") and disadvantages ("superficiality" and "subjectivity"). At IHS a comprehensive evaluation model was developed for the evaluation of university research (and university-related research). This model proposes a dual evaluation mode by combining equally a monitoring system and a system of *ex post* evaluations. In accordance with evaluation results, "research points" may be accredited to university departments.

Literature

Bell, *Daniel (1999) (first published in 1973). The Coming of Post-Industrial Society. A Venture in Social Forecasting. New York: Basic Books (ISBN 0-465-09713-8).*

Campbell, *David F. J. (1999). Evaluation universitärer Forschung. Entwicklungstrends und neue Strategiemuster für wissenschaftsbasierte Gesellschaften.* SWS-Rundschau *39 (4), 363-383 (swsr@aon.at).*

Campbell, *David F. J. / Bernhard* Felderer *(1997). Evaluating Academic Research in Germany: Patterns and Policies. Vienna / Institute for Advanced Studies (IHS): Political Science Series No. 48 (library@ihs.ac.at).*

Campbell, *David F. J. / Bernhard* Felderer *(1999). Empfehlungen zur Evaluation universitärer und ausseruniversitärer Forschung in Österreich. Vienna / Institute for Advanced Studies (IHS): Political Science Series No. 66 (library@ihs.ac.at).*

Felderer, *Bernhard* / *David F. J.* Campbell *(1994). Forschungsfinanzierung in Europa. Trends, Modelle, Empfehlungen für Österreich. Vienna: Manz (ISBN 3-214-05998-X).*

Gibbons, *Michael* / *Camille* Limoges / *Helga* Nowotny / *Simon* Schwartzman / *Peter* Scott / *Martin* Trow *(1994). The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. London: Sage (ISBN 0-8039-7794-8).*

VSNU – Association of Universities in the Netherlands (1994). Protocol 1994, Utrecht.

VSNU – Association of Universities in the Netherlands *(1998). Protocol 1998, Utrecht.*

Author reference

David Campbell, Institute for Advanced Studies (IHS), Department of Political Science, Stumpergasse 56,
A-1060 Vienna, Austria
Tel: ++43-1/59991-173
Fax: ++43-1/59991-171
E-Mail:David.Campbell@ihs.ac.at
Internet: http://www.ihs.ac.at

Evaluation at the Austrian Academy of Sciences

Armin Scrinzi

During the last five years the research units of the Austrian Academy of Sciences (AAS) were systematically evaluated by external peers. Results and consequences of the evaluation are being prepared for publication. This article summarises experiences and illustrates the consequences of this first comprehensive evaluation of a entire research institution in Austria.

I. Motivation and purpose of the evaluation

Anticipating the increased public demand for transparency in using research funds, the AAS published in 1996 a programmatic brochure intended for the general public with the title "Mittelfristiges Forschungsprogramm 1996 - 2000" ("Intermediate Time Scale Research Program 1996 - 2000", [1]), which gave an overview of the activities of the Academy's research units and an outlook on their future research.

The brochure was aimed at initiating a model process for the public discussion of research programs and it included the project of external evaluation as one key element in this process. The evaluation of the research units of the AAS was to serve as an advisory instrument for strategic research decisions of the AAS. Other than that, it should also provide objective feedback to the units about the quality of their work and fulfil the perceived duty of the AAS to be accountable to the public. The latter goal was then, as it is now, an important argument in the increased competition for public research funds. On a more general level, the AAS hopes to contribute to a „culture of evaluation" in Austrian research and by that to stimulate conceptual research politics, the absence of which is frequently deplored.

II. Specifics of the research in the AAS

Publicly funded non-university research in Austria is very diverse in terms of scale, subjects, structures, and tasks. It constitutes a comparatively small sector of Austrian research, but there seems to be general agreement that it should be developed further, for which some form of evaluation would provide a rational basis. In view of the large diversity of the research it is obvious that the procedures of evaluation must be adapted to the existing structures. By taking the initiative in organising its own external evaluation the AAS had the freedom to adjust the procedures to its specific needs. The brief overview of the AAS's research is given here to put the adopted procedures into the right perspective.

Diversity of research subjects

The AAS is the largest non-university institution of basic research in Austria. Its research covers a very broad range of subjects in science and in the humanities. These include subjects as diverse as archaeology (Forschungsstelle Archäologie), molecular biology (Institut für Molekularbiologie), space science (Institut für Weltraumforschung), editions (Corpus Scriptorum Ecclesiasticorum Latinorum, diaries of Arthur Schnitzler), comparative behavioural science (Konrad Lorenz Institut für Vergleichende Verhaltensforschung), Asian studies (Institut für Asienforschung) sociology and political sciences (Institut für Technikfolgen Abschätzung, Forschungsstelle für Institutionellen Wandel in Europa) etc.

Diverse research tasks

Related to the various subjects AAS research also has different tasks. Part of the research is similar to typical university research, i.e. it can be structured in

short- or mid-term research projects with the duration of a few years. Justification to conduct such research outside the universities is by complementarity, i.e. the AAS strives to open up new fields of research for Austria that are not or not yet represented at the universities. A typical case is the Institute for Demography, which is - because of the total absence demographic institutes at the universities - the only institute of its kind in Austria, or strongly interdisciplinary research, as for example the Institute of Technology Assessment.

As a national academy of sciences, the AAS has to fulfil tasks of national interest, such as the compilation of lexica and data bases and the operation of archives that are of special importance in the national context. Examples of such project are the „Wörterbuch der bairischen Mundarten in Österreich“ and the corresponding data base, or a Lexicon of Austrian Music, and similar tasks in the history of art, archaeology on Austrian territory or medieval history. These projects are by their nature long term and cannot be adequately evaluated by the methods and criteria of mainstream science, especially if these criteria were to be limited to the number of publications, conference contributions, output of students, or number of patents. Still they must be subject to rigorous scientific standards, which are, in general, well understood within the scientific communities relying on them.

Another scientific task of the AAS are long term international projects, where formally and in practice a national institution guarantees reliability and constancy of contracts. The traditional and highly reputable Austrian research in Asia Minor including the excavations in Ephesos falls into this category. Similar to archives and editions, these projects must be judged by general scientific criteria as well as by their national interest and their role in providing infrastructure for the whole field of research.

Research structures

The diversity of subjects and tasks is matched by equally varied organisational forms. The AAS knows three distinct types of research units: Institut, Forschungsstelle (research group) and Kommission.

An "Institut" is comparable to a small university department or major independent university research group.

A "Forschungsstelle" is a transient organisational form which after a period of 6 years must either be made permanent in the form of an Institut or else must be dissolved. Both forms have an international scientific advisory board.

A "Kommission" is primarily the association of scientists and scholars for the pursuit of a specific project or purpose with no or very few employees. In spite of this seemingly ephemeral character of a Commission some of the longest ranging and most reputable research of the AAS is executed by Commissions. For example, editions are typically organised in Commissions.

Small scale research

Finally an important point of AAS research is that on average it is small scale. A total of about 50 research units shares an operational budget of approximately ATS 250 mil (Euro 18 mil). Cost of infrastructure as office space, telephone and heating is paid in addition and is not included in that figure. As mentioned above, a few of the institutes are university size, but the majority of Commissions shares a small part of the budget.

III. Evaluation principles and procedures

The principles and the purpose of the evaluations at the AAS were published the "Intermediate Time Scale Research Program 1996-2000" [1] and were restated

in a second brochure "Evaluationsergebnisse I" ("Results of the Evaluation I", [2]).

Principles

Key principles are the following:

- Evaluation serves as an advisory instrument for AAS research planning and it should help to identify future research options. Its primary interest lies in the future of the research field.
- Several related research units are assigned to a „research area“. The subject of evaluation is a whole research area within the AAS, as well as the separate research units. Researchers are not subject to the evaluation as individuals.
- Evaluation defines the quality standards for a research field and it is to give feedback to the individual research units about the quality of their work.
- The evaluation is performed by a group of external (mainly international) peers. The AAS takes no influence on the composition of the group other than naming an independent expert as chief evaluator. The chief evaluator selects his evaluation group autonomously.

Procedures

The following procedures were applied to all evaluations:

(1) The individual research units publish research programs.

(2) Several units with related research are joined to form a „research area“ that will be evaluated by the same peer group.

(3) The AAS designate a chief evaluator for the research area. The chief evaluator forms an evaluation group with typically 4 to 6 members.

(4) The evaluation group is supplied with information on structure, budget, main publications, the research unit‘s answers to a questionnaire and a list of questions of the AAS to the evaluation group. The questions to the research unit are worked out in collaboration between the AAS administration and the chief evaluator.

(5) The AAS presidency meets the evaluation group for information about special questions concerning the AAS research area. An evaluation visit including talks with the heads of the units, with staff and other personal and visit of offices, labs etc. follows.

(6) The evaluation group supplies a preliminary report to the AAS administration and to the research units. The units are invited to comment on the report. The comments are brought to the attention of the evaluation group before a final report is written.

For part of the evaluations one or both of the following stages were added:

(1) A symposium on the research area or on one of its research units is arranged. The purpose of the symposium is the public discussion of research programs and future perspectives of research on a national scale.

(2) After receiving the final report the research units are invited to comment on procedural questions and on the implementation of recommendations in the report.

For a more complete description of the procedure including details on the written materials supplied by the research units the reader is referred to the brochure "Evaluationsergebnisse I" (in German, [2]).

IV. Consequences of the evaluation

Several important decisions in AAS research were either directly triggered or strongly influenced by the evaluation reports. The changes can be grouped into changes in research orientation, personnel, infrastructure, and budget, and, most importantly, research structures. The overview of consequences given here remains incomplete, since several of the reports are fairly recent (fall 1999) and not all possible consequences have been implemented yet.

New research

The orientation of research and the implementation of corresponding recommendations ultimately remains the responsibility of the research units. An example for the reorientation of the research in an existing institute is the Erich Schmidt Institute for Material Sciences (formerly Solid State Physics) where under a new director a clear emphasis was put on Material Sciences in accordance with the institute's new name. The Acoustics Research Institute has reorganised its research following recommendations in the evaluation report. Several institutes and commissions have taken up projects following suggestions or encouragement by evaluation groups.

Personnel, budget, and infrastructure

Numerous recommendations concerned personal, budget, and infrastructure. As may be expected, only a certain part of such recommendations can be realised because of budgetary, topographic or structural limitations. Until now a total of 14 new science and administration jobs were created in accordance with the recommendations. Several units have been relocated to more adequate locations and computers and other infrastructure were extended. In general, it is difficult to judge the exact contribution of evaluation to such decisions, which are a natural part of everyday administration, but certainly several problems have received additional attention in consequence of the evaluation.

Research structures

The most visible and most debated consequences of an evaluation are the foundation of new institutes, fusion of existing units into a single larger one, but also the closure of research units. Several such decisions were made as direct or indirect consequences of the evaluation reports.

The "Forschungsstelle für Sozioökonomie" was closed. It was replaced by a "Forschungstelle" for "Institutional Change in Europe" with new tasks, a new director, and partially new personnel.

Although not recommended in the evaluation report, but as the consequence of structural problems pointed out therein, it was decided to close down one externally located department of the Institute of Limnology by the end of 2003.

The Institute for Information Processing was dissolved. One of its former departments formed the basis for the newly founded "Institute for Discrete Mathematics".

Several "Kommissionen" were closed.

V. Acceptance by the research units

Initially the project of evaluation was met with scepticism, but in course of time objections appeared to diminish. In order to form a more objective picture of the process and to offer a forum for criticism of the procedures, research units whose evaluation was finalised were invited to comment on the evaluation process as a whole. Through this feedback the administration became aware of several problems of the evaluation that are discussed in the next section. It is fair to say that the majority of the research units had no fundamental objections against the evalua-

tion. Several units explicitly stated a positive, stimulating effect of the evaluation process itself on internal communication and awareness of the unit's common mission. In most cases the site visits of the evaluation groups were perceived as benevolent and the talks with the evaluation group sometimes extended into a scientific exchange.

It was criticised in several cases that in spite of the friendly atmosphere during the visits, fundamental points of the research could not be communicated and were therefore not adequately reflected in the final reports. Other criticism referred to insufficient information by the administration and to the extensive effort and paperwork for the preparation of the evaluation.

VI. Problems of the evaluation

Research quality criteria

A fundamental problem of any evaluation is the contradiction between the administration's need for simple „objective" criteria for the quality (and quantity) of research output and the necessarily complex and sometime elusive nature of creative research. While in several major branches of science a large international community has accepted standard quantitative criteria, no such criteria are available in much of the humanities and part of social sciences. This problem remains important for the advisory type evaluation performed at the AAS, but it is mitigated by the replacement of the one- or multidimensional grading of the research units by specific recommendations for research decisions.

Selection of peers

As in other peer review systems, the problem of establishing objective criteria resurfaces as the problem of objectively selecting the peer group. The solution found for the AAS - nomination of a single, independent, highly qualified chief evaluator, who is solely responsible for the further selection of the group members - was meant to make the group immune to interventions from inside the AAS. The strategy was successful in that respect, but in some cases the danger of obtaining a one-sided selection of peers arose. For future evaluations, the right of the research units to comment on the composition of the evaluation group before its final formation should be included, as well as, in the extreme, the right to object to individual members of the group.

Definition of the research areas

Similar as in any evaluation, the AAS faced the problem assigning independent but related research units to research areas for common evaluation by a single peer group. The assignment was performed by the AAS administration. One of the originally intended areas turned out to be too heterogeneous for finding a single evaluation group and it had to be split up. In other cases dissatisfaction with the assignment was aired by research units. It should be emphasised that the decision of grouping research units into research areas is not purely pragmatic. Its purpose is to avoid judgement by only very narrow criteria of a single research field, as it usually would be preferred by a single research unit. Research decisions need to go beyond narrow disciplinary boundaries and reports that take a more general point of view are of greater use. Again, in future evaluations improved communication with the research units prior to the final assignment to research areas should help to reduce this problem.

Technical problems

The remaining problems were of a more technical nature, mostly owing to the of prior experience with evaluations both, on the sides of the administration and of the research units. Such problems include

misunderstandings about the purpose and possible consequences of the evaluation, insecurity about the procedures or unclear definition of the task of the peers. More systematic information and increasing experience on all sides should be able to solve that type of problems.

Integration of evaluation into the decision process

Evaluation as an organisationally separate, external procedure of quality assessment is a new phenomenon. Its connection to the regular administrative procedures still has to find its final form. The practice in the AAS has been to make rather non-formalised use of the evaluation reports, where recommendations were implemented according to the possibilities and general intentions of the Academy‘s research politics. An important element of the evaluation is the publication of its results, which subjects the implementation (or non-implementation) of recommendations to public discussion. The author considers it desirable to parallel the indirect public mechanism of discussion by a formal internal procedure, which should lead, in given time, to a summary statement about implemented and rejected recommendations of the evaluation.

Symposia

The symposia, which had the ambitious goal to stimulate nation-wide programmatic discussion of certain research areas, did not succeed in this. They hardly had any identifiable impact beyond being a showcase for a given research unit. One reason is that in some research areas on a national level there are very few or no other institutions outside the AAS with sufficiently closely related research interests, rendering discussion within an Austrian framework meaningless. Another problem may be that in a symposium, where one institution was subject to evaluation and others not, a common point of reference for the discussion of research programs is missing.

From a general point of view, the public discussion of the quality and purpose of research remains highly desirable. Evaluation, if established nation-wide, may serve as a natural starting point for such a discussion.

VII. Cost of the evaluation

One significant part of the cost of evaluation is the time and effort spent in the research units on preparation for the evaluation. That kind of cost is difficult to quantify, but it has led to complaints from several research units. The experience of one of the Academy‘s evaluators, who had himself undergone three evaluations during a single year (by different institutions), should warn us against the perils of an "evalumania". With a moderate use of the instrument of evaluation and with increasing routine on the side of the researchers this part of the effort should remain acceptable.

The financial cost of a typical evaluation with six peers evaluating six research units during four days was about ATS 400 000.- (Euro 30 000.-). More than half of that cost are fees paid to the peers. Having in mind the common practice of uncompensated peer reviewing of journal articles, the compensation for the evaluation may seem unusual. However, the much larger effort for an evaluation, the considerable responsibility connected with it, and the need to attract scholars of highest possible competence and reputation fully justify such a compensation. Moreover the cost of the evaluation must be compared to the total research budget for the period between repeated evaluations, where it amounts to much less than half a percent.

VII. Future evaluations and conclusions

The generally positive experience with evaluations at the AAS is reflected in the Academy‘s decision to repeat the evaluations in seven year intervals. At present, no fundamental changes of the procedures are planned but several adjustments as outlined above will be made.

Evaluation as conducted at the AAS, which involves time-consuming and expensive site visits, does not appear to be feasible at much higher frequency than seven or six year intervals. In the author‘s opinion, the peer review evaluation could be advantageously supplemented by shorter term evaluations using bibliometrical methods that should be related to the peer review reports. Criteria relevant for the individual research units could be proposed by and discussed with the evaluation group and might comprise data like increase or decrease of publication activity, conclusion and acquisition of new research projects, the development of collaborations, the pace of progress of long-term projects etc. The purpose of such controlling would be to monitor the success of new strategies or projects and to have an early indication of possible problems.

The decision to continue evaluations implies that evaluations are seen as advantageous for the Academy‘s research. Although this general impression about evaluations is naturally shared by the evaluation community and also by the legislator, an empirical prove is difficult to conduct. At the meeting, where this presentation was originally given, it was pointed out on the example of the British research output that at present no clear signature of an impact of the harsh British evaluation system on bibliometrical parameters can be detected. Chances to prove effects on more evasive concepts such as the quality of research and, connected with that, its efficiency seem even less.

On the other side, there is the experience of stimulating discussions with the external evaluation groups and implemented recommendations of the evaluation reports, which prove that the approach to evaluation chosen at the AAS can be put to use. Whether the principles applied at the AAS can be carried over to a larger number of research institutions remains to be explored.

Beyond the impact on science itself, only the future will show whether evaluation - in which form ever - can satisfy the legitimate demand for accountability of publicly funded research to the public.

Literature

"Mittelfristiges Forschungsprogramm 1996-2000" ,edited by the Presidency of the Austrian Academy of Sciences, Vienna, 1996

"Evaluationsergebnisse I" (ibid, 1998)

Both brochures (in German) can be obtained from the Austrian Academy of Sciences, Dr. Ignaz Seipel-Platz 2, A-1010 Vienna, Austria

Author reference

Armin Scrinzi works in physics research at the Vienna University of Technology. From July 1998 through June 1999 he was in charge of the administration of the evaluations at the AAS.
Einsiedlerplatz 2/10, A-1050 Vienna
E-mail: scrinzi@tuwien.ac.at

Evaluation of Biomedical Research in Austria

Bernd R. Binder

Evaluation in general

In every society evaluation of the performance of members or groups within that society is performed almost on a continuous basis. It is either the free market where the quality of the products produced is "evaluated" by the success of the business at large (from the butcher's shop to industry) or it is the evaluation by peers in a "non-free market" situation such as within the family, the school, a non market oriented group, a religion, or a whole society in an authoritarian system. Without evaluation and resulting consequences (benefits or punishments) every individual or group will tend towards a lifestyle minimising the necessary efforts to obtain all material and non-material goods to satisfy the needs of that particular individual or the group (such a society could be called "free running" or "free floating" society). Consequently in such a "free running" society, performance of any individual or group will solely depend on the specific personality, the demands of that person for own satisfaction (money, influence, recognition) [1]. Free running groups that neither produce goods for the free market nor have - at least for specific parts of their lives - peers, exist within every society. The best example is - besides groups with no demands ("flower people") – the peers themselves. The danger of a "free running" group is obvious: If the members of such a group are not highly motivated and are not - as individuals - up for high standards (at least what by high standards is understood by that specific society), this group will sooner or later adapt to minimal demands. Usually - at least up to recently in academia - a person became only a peer when his or her performance was outstanding and in most cases the person did not change in his or her personality pattern upon becoming a peer. Therefore peers as a "free running" group without evaluation by the free market or by other peers retained their high standards and did in general not deteriorate.

Evaluation of Universities

In the European, specifically in the Austrian university education system, the free market is not applicable. Students do not pay for their education and therefore the quality of teaching and overall education does not give a feedback to the university or to the individuals involved in education in form of financial or promotional benefit. Like other educational systems universities are classical peer evaluated systems: Students are evaluated by their teachers and, in a classical hierarchical system, every level of the system is peer evaluated by members of the upper one and up to the top; as a consequence the relative proportion between being peer evaluated and being oneself a "free running" peer shifts towards being only a free running peer. Again, as long as the peers fulfil high standards, the whole system will function to the satisfaction of the society.

The problem starts only whenever an appreciable percentage of the peers become peers not because of their excellent performance but for other reasons or - not to be neglected - the expectation of the peers deviates from the demands of the society. That all applies to the educational part of universities and also to their research activities. It is, however, important to note that, especially for research, universities are not only embedded in the local community but also in the international scientific community. Thereby a deviation can also occur in so far as local

[1] Arnold J. Toynbee: A study of History, Abridgement of volumes I – VI by D. C. Somervell, Oxford University Press, London, 1950

research expectations might grossly differ from international standards. Such deviation from local to international standards occurs specifically in a closed or pseudo-closed society such as research in the former Eastern countries or research in societies where spending for research differs from international standards and therefore local research cannot compete with the international mainly due to lack of resources (e.g. percentage of the GNP spent for research significantly below average of a comparable scientific community).

Evaluation of research in Austria

Until World War II there was more or less an intact peer system at Austrian Universities. Peers fulfilled high standards and expectations of the local and international community (scientific and non-scientific) coincided with those of the peers. After World War II the peers were depleted due to the loss of Jewish peers making up a rather high percentage of the peers before 1938 (e.g. the medical faculty lost about 50% of their academic teachers in 1938, and only a few returned after the war[2]). Furthermore, as compared to research in the US, financial resources were extremely limited resulting soon in a deficit to introduce new and expensive devices and technologies. As a result research standards deviated between those applicable and therefore applied in Austria and internationally, especially in the US. In addition, after 1968 and followed in the early 70s by the introduction of the new University organisational law (UOG 73), the view of the society, how universities should be functioning, became increasingly different from what at that time ruling peers expected. As a consequence, the policy for recruiting new peers changed and the group of peers became more and more non-homogeneous. This resulted in a malfunction of the peer system, because the peers themselves, as a "free running" group being under no external pressure were not able to set uniform standards anymore. This resulted in vast differences between different clinics and different basic science institutions in e.g. standards for promotion of personnel or the way science was conducted and communicated to the scientific community. In turn the scientific output differed dramatically between different institutions at the same faculty. In addition at that time, distribution of resources was not dependent so much on scientific or educational achievements but did depend on personal contacts to peers within the university or the administration.

Evaluation of biomedical research output by the Vienna Medical Association

This situation – specifically the fact that distribution of resources was not obviously linked to scientific achievements - resulted in the late 80s in initiatives within the Medical Faculty of the University of Vienna to evaluate biomedical output in a more objective way with the final goal to link scientific output to the distribution of resources. For this purpose, the bibliographic method was chosen and in 1987 the Vienna Medical Association performed the first evaluation of biomedical research output which from that time on was conducted on a yearly basis. These public reports are presented at an annual meeting of the society.

The bibliographic method for evaluation of research

The bibliographic method is based on the impact factors of journals allocated annually to the relevant journals in certain fields, most prominently biomedicine. The impact factor of a journal is the average frequency of citations per year of an article published

[2] W Schütz, K Holubar, W Druml (1998) Editorial: on the 60th Anniversary of the Dismissal of Jewish Faculty Members from the Vienna Medical School. Wien Klinische Wochenschrift 110, 113-114

in that particular journal and depends on the scientific quality of the journal and its area of (international) distribution. The bibliographic method then allots to an article published by an institution (or an individual) points according to the impact factor of the journal in which that article was published. In general, the more prestigious a journal, the higher the impact factor of that journal and the more difficult to get an article accepted in such a journal and – usually - the better the quality of that article. Such a correlation holds true not on an individual article by article basis but only statistically. Articles published in a high impact journal are better on average and are more frequently cited than articles published in journals with low or even missing impact factors. Besides the need to be cautious when judging the quality of an individual article only by the impact of the journal in which the article was published, trends and fashion in science also have an influence on the impact of a specific field and therefore on journals in a particular field. Therefore "trendy" fields achieve usually higher impact points than articles published in fields not following the main stream of research.

The bottom-line, however, is that international visibility of research of a specific individual, institution or country is pretty well reflected by the bibliographic method that also allows easy comparison between individuals, institutions or regions. One might win the Nobel Prize with an article published in a non-impact journal, but this is highly unlikely. This all justifies the use of the bibliographic method to evaluate scientific output.

When evaluating institutions according to the bibliographic method, the problem of allocating publications to an institution exists. As long as an individual publication originates only from one institution, allocation of that particular publication is easy. In modern biomedical science, however, collaborations are indispensable and resulting joint publications have to be allocated. One could use a percentile allocation of impact points to an individual institution that would likely be the most fair way. Such a distribution, however, makes computer-assisted allocation almost impossible. We therefore have chosen to allocate a publication only to one institution, namely the one of the corresponding author. It is likely that the main resources necessary to perform the work leading to that specific publication originates form the institution of the corresponding author. In addition it is highly unlikely that collaborations are only "one-way" resulting in all publications originating from such collaboration being published only from one institution. On average, allocating the points from a publication only to the institution of the corresponding author would not be unfair and, on the other hand, easy to administer.

Based on these considerations, the bibliographic method in the modifications as outlined above was used from 1987 on to evaluate Viennese and Austrian biomedical output. This bibliographic method, however, does not allow comparing different fields of biomedical science, because of the different "impact" of the particular field. Reasons for such differences are given above.

Ranking according to" top", "standard" and "other" journals

From 1992 on an additional parameter was introduced: To be able to account for differences between different fields in biomedical science, which – as outlined above – would achieve different impact points, the term "top" journal and "standard" journal was introduced. A top journal was defined as a journal ranking within the top 10% or – for purposes of promotion to become a "Docent" – top 20% of journals listed in a specific field of biomedical sciences. This list is published annually by the Institute for Scientific Information in the USA, the same institution

that also publishes the impact factor of the different journals. A standard journal is defined as a journal ranking between the top 10% (or 20%) and 60% of the journals listed in the field. Journals ranking below 60% are neglected. This categorising of journals into groups allows comparison between different fields of the biomedical science and can therefore be applied as standard for promotion or distribution of funds from e.g. the Faculty to different clinics and basic science institutes, respectively. This method could also be applied for an easy comparison between the scientific output of different universities, cities and countries. Therefore, from 1992 on Munich, Heidelberg, Göttingen, Basel, Zurich and Lausanne were also evaluated as well as Germany, Switzerland, Denmark and Belgium and compared with respect to their scientific output in biomedical sciences to Vienna, Graz, Innsbruck and total Austria.

Results of the bibliographic evaluation of biomedical research output in Austria from 1987 to 1998

During that 12 years period and in more detail during the period from 1992 till now, biomedical research output from Austria increased significantly. In Fig. 1 and 2 the total biomedical scientific output over time is given and compared to the scientific output in selected countries in Europe. To account for differences in the specific countries, data are given per capita (Fig. 1) per gross national product (GNP), and per money spent for research and university research, respectively (Fig.2). As can be seen, the total biomedical scientific output in Austria as well as the top scientific output (number of 10% top papers published) is better than that originating from Germany but much worse than the ones originating from the other countries used for comparison (Switzerland, Belgium and Denmark). When not only population is used for comparison but also resources a similar situation is seen: Using the parameter biomedical publications per GNP, Austria still ranks before Germany and approaches Denmark. Publications in biomedical sciences related to money spent for a university interestingly increases the gap again and now the curves for Austria and Germany are actually superimposable (Fig. 2). This could be interpreted that in Austria funds attributed to “Universities” are actually not spent for research but rather education and relating research to funding institutions would be better suited.

When one uses the same algorithm for comparison of the three Medical Faculties in Austria, Graz, Innsbruck and Vienna for the clinical part and for preclinical basic science institutions data as shown in Fig. 3 and 4 are obtained. As the total scientific output not only depends on the quality of the research personnel but also on their number, a comparison of the research output for the three Medical Faculties per academic position is given (Fig. 3). From this figure it can clearly be seen that efficiency of scientific output (published impact per academic position) was best for Innsbruck for a long time and only recently was overtaken by Vienna. The Innsbruck preclinical institutes still remained more efficient than any other group in one of the Austrian Medical Faculties (Fig. 4). While the increase in scientific output above the long lasting trend for Vienna 1993 can partially be attributed to the fact that the clinics had moved to the new general hospital with the respective outstanding facilities there, at least two additional explanations could also be offered. Both are linked to evaluation: One is the fact that at the same time the new rule in Vienna for promotion to become a “Docent”, the major step in the university career, became effective. These new rules demand a at least 11 “standard” publications and at least 1 “top” publication. Since an appreciable part of University research is done by academic personnel aiming to become a “Docent” such a change should have a major impact

on the overall scientific performance of a faculty: This change in promotion rules was most effective in Vienna at that time. The Medical Faculty of Innsbruck had already introduced a different, but also demanding rule earlier, while the Medical Faculty of Graz has not done so yet (Fig. 4). The other explanation might be the fact that evaluation was actually performed for the Vienna Medical faculty from 1987 on, although a link between scientific performance and resource allocation was introduced only recently.

Effect of the evaluation of the scientific output in the Medical Faculty of Vienna

When one follows the scientific output for individual institutions in Vienna, an interesting phenomenon is appreciated: There seems to be no change in the performance trend for institutions with very low or even no scientific output over time, regardless whether data for the total scientific output or the scientific output per academic position are used. This indicates that these institutions are not influenced by the fact that their low performance is publicly known. On the other hand, when the institutions with the highest scientific output either total or per academic position are analysed, these institutions all have an upward trend of their performance. At the public evaluation meetings as expected no representative of the low performing institutions was present, while in general the heads of the high performing departments were always present. This indicates that a challenge or a competition leads to better performance only for those who can meet that challenge already. For those who are overwhelmed by the competition, the competition as such does not increase their performance. This pattern clearly also follows the theorem put forward by Arnold J. Toynbee in his "Study of History" that a society with either no challenge or facing an overwhelming challenge will deteriorate and not prosper (Challenge and Response[3]).

Feedback between scientific output and resource allocation

In addition to the benefits of the mere fact of an evaluation for institutions within the competition for the front places, a link between resource allocation and performance might also boost the performance of institutions with low performance which are therefore not challenged by the competition as such. In 1997 the Medical Faculty in Vienna adopted a modified version of the bibliographic method of evaluation as one of the decisive factors to allocate funds. This policy is now effective for the third year and it is simply too early to judge whether this link between performance and fund allocation has also changed the performance pattern of the institutions with a low scientific output. Perhaps evasion mechanism such as "publishing in non-impact factor journals", "being in a field where the impact factor is not important", "having no time for research because of overwhelming teaching responsibilities" are still too effective and scientific output becomes not so important that everything is done to perform better in that particular respect.

Abstract

Evaluation of biomedical research in Austria was initiated by the Viennese Medical Association in 1987 and from that time on continued from that time on an on an annual basis. A modified bibliographic method is used and the impact points from an article published are allocated to the institution of the corresponding author. Results from this evaluation are published at an open meeting of the Viennese Medi-

[3] Arnold J. Toynbee: A study of History, Abridgement of volumes I – VI by D. C. Somervell, Oxford University Press, London, 1950

cal Society and in the meantime are used by the Austrian Government and – in a further modified version – by the Medical Faculty of the University of Vienna to allocate funds. Total scientific output in Austria increased significantly during the evaluation period and is better compared to Germany but inferior to the one from Switzerland, Belgium and Denmark. Within Austria, the impact of evaluation seems to be most prominent for Viennese biomedical science; relative scientific output (published impact per academic position) originating from the Vienna Medical Faculty overtook that from Innsbruck being first for several years previously. Evaluation as such seems to be beneficial only for institutions competing for front places but not for institutions having low or no published impact. For the latter, a stronger link between scientific performance and allocation of funds might also improve their scientific performance.

The figures of this article can be seen on the following two pages.

Author reference

Bernd R. Binder, Professor of Physiology and Head of the Department of Vascular Biology and Thrombosis Research at the Medical Faculty of the University of Vienna, Schwarzspanierstr. 17, A-1090 Vienna
Tel. ++43-1/4277-62501
E-mail: bernd.binder@univie.ac.at

Figures

Figure 1: Number of biomedical scientific papers published in the years indicated for different countries. Bars: Total per year, Line: per 10 million population.

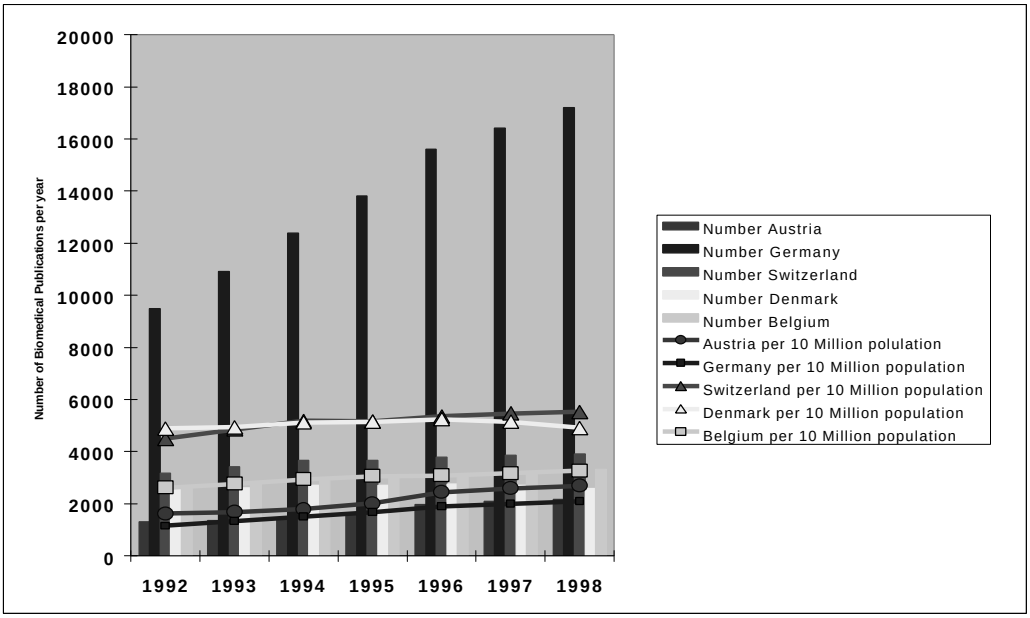


Figure 2: Number of biomedical scientific papers published in the years indicated for different countries. Bars: Number per 1.000 Million USD GNP; Lines: Number per 10 Million USD spent for research and universities, respectively.

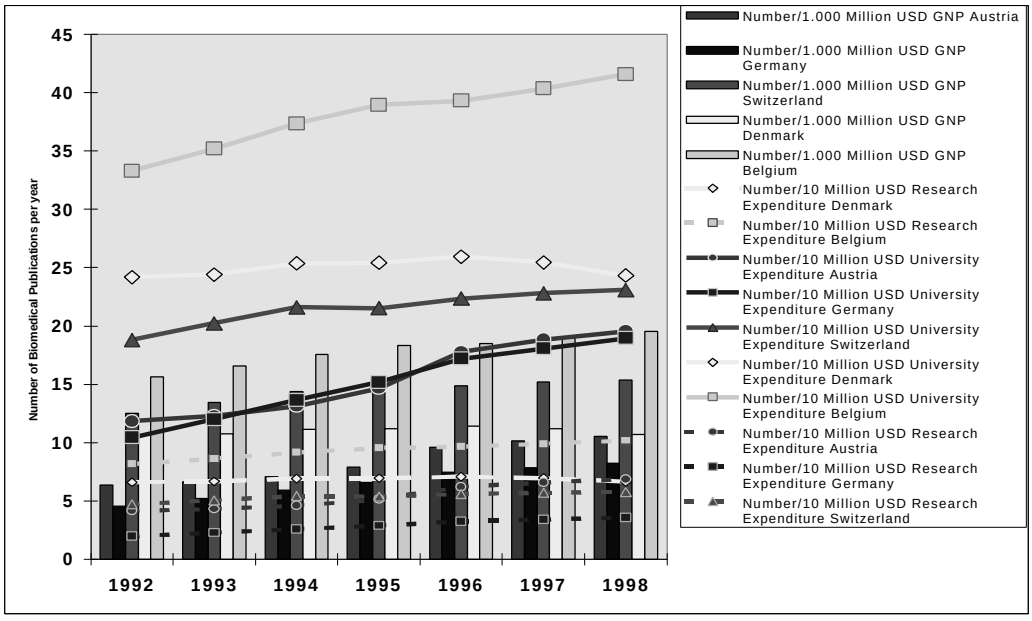


Figure 3: Biomedical publications (bars) and biomedical impact points (lines) per academic position within the medical faculties of Vienna, Graz and Innsbruck, respectively

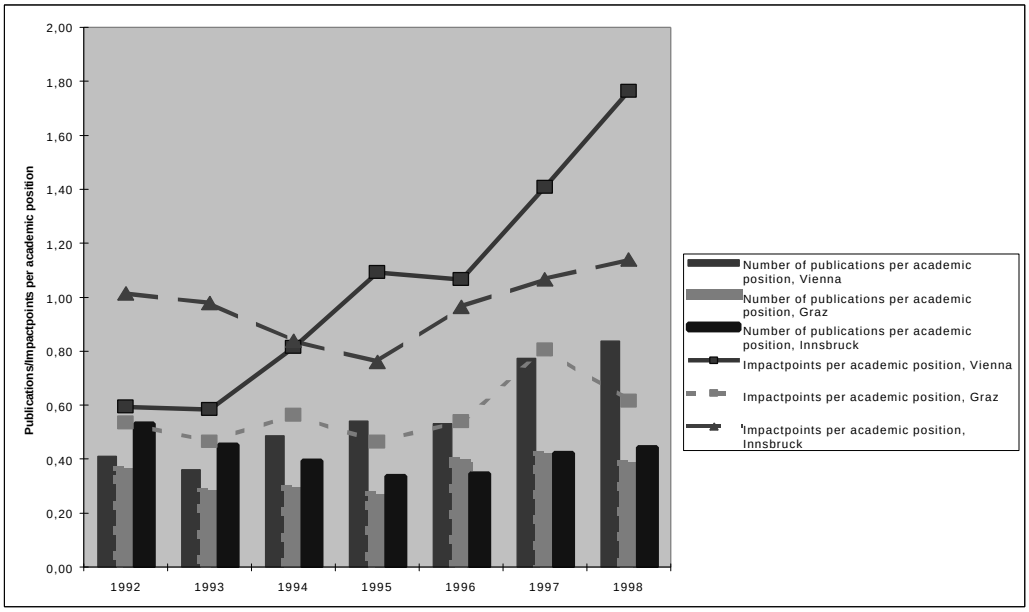
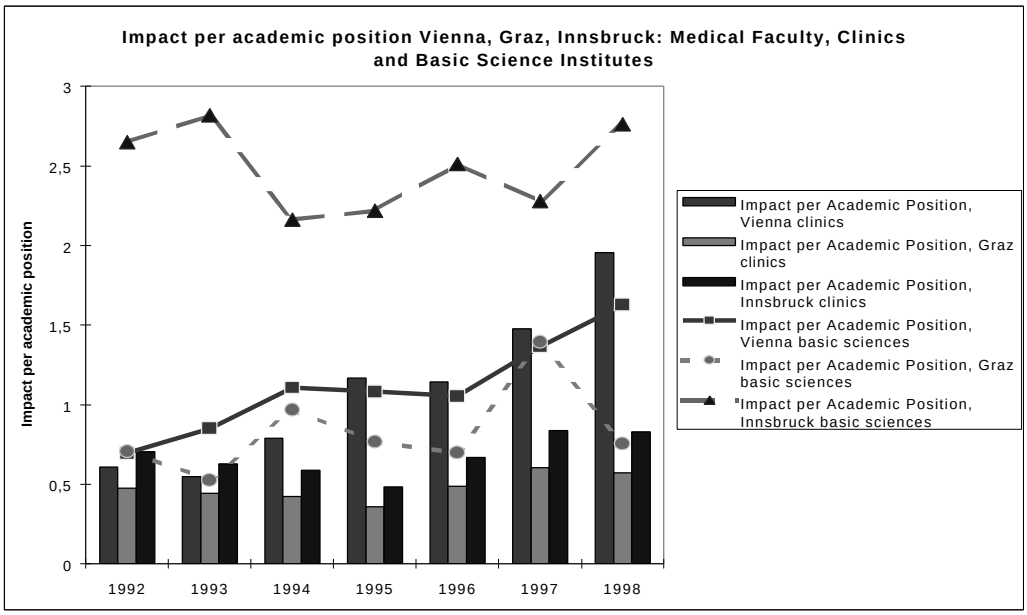


Figure 4: Biomedical impact per academic position per year during the period 1992 to 1998 for clinics and basic science institutions at the medical faculties in Vienna, Graz and Innsbruck, respectively.



Berichte

Evaluierungsseminar in Karlsruhe am 16./17. 3.2000

Evaluierung wettbewerblich organisierter Multi-Akteur-, Multi-Maßnahmen-Initiativen

In der Technologiepolitik und auch in der innovationsorientierten Regionalpolitik tauchen zunehmend Programme auf, in deren Rahmen mehrere Arten von öffentlicher Förderung (Forschungsprojekte, Vernetzung, Risikokapital, Institutionenaufbau) mehreren Kategorien von Akteuren (Unternehmen, Start-ups, Universitäten, öffentliche Agenturen etc.) zukommen. Diese Initiativen werden unter einem gemeinsamen Dach entwickelt, verkündet und umgesetzt; in der Regel von einem Projektträger betreut. Ausschreibungen sowohl zur Ermittlung des Projektträgers als auch zur Findung der Projekte sind dabei die wichtigste Verfahrensform. In Deutschland stellt etwa das BioRegio-Programm eine Initiative dar, in der alle diese Elemente stark entwickelt vorkommen.

Für die Evaluierungsseite stellen derartige große, multiinstrumentelle und viele Akteure einbeziehende Programme eine Herausforderung dar:

- Wettbewerbe haben bei der ex ante-Auswahl eine Reihe von Vorteilen, bergen aber auch einige Schwächen / Gefahren in sich (etwa Kostenprobleme, Informationsasymmetrien, ambivalente Signale in das Akteursfeld, Verliererproblematik, usw.)
- Durch die sehr weiten und auch oftmals weichen Zielsetzungen der Programme sowie das Vorhandensein oft ganzer Maßnahmenbündel ergeben sich für das Monitoring und die ex post-Evaluierungen Schwierigkeiten in der Abgrenzung, der Festsetzung von Indikatoren und bei der Messbarkeit von Ergebnissen. Vorhandene Indikatoren / Instrumente sind weiterzuentwikkeln.
- Dabei zeigt sich einmal mehr, dass eine ordentliche qualitative Evaluierung (angereichert um den einen oder anderen quantitativen Indikator) in der Regel zu besseren Ergebnissen führt als eine gnadenlose Quantifizierung.
- Evaluatoren geraten gelegentlich in Konflikte dadurch, dass sie auch in die Rolle von Programmcoaches hinwachsen (in Deutschland in einigen Fällen) und daraus ein Rollenproblem entstehen kann.
- Programmdesign und auch Evaluierungspraxis hängt auch von neuen Formen der Schaffung von Wissen in Innovationssystemen ab.
- Schließlich stellt sich die Frage nach der politischen Gestaltbarkeit derartiger Settings: kann man Innovationssysteme "miterschaffen", indem man besonders gute Förderanreize etc. setzt? - und schließlich was kann dabei die Rolle von Evaluierungen sein?

Zu diesen Fragen wurde unter organisatorischer Leitung des ISI Karlsruhe ein Workshop abgehalten, bei dem konzeptive und empirische Arbeitsergebnisse vorgestellt wurden; namentlich

- eine Abwägung von Vor- und Nachteilen von Wettbewerbsverfahren mit Schwerpunkt auf der Ausgestaltung und Funktion der ex ante-Bewertung (Sturn, Joanneum Research, Wien),
- eine Übersicht zu laufenden Multi-Akteurs-, Multi-Maßnahmen-Programmen in Österreich vor dem Hintergrund spezifischer Muster in der Technologiepolitik (Stampfer, TiG, Wien),

- eine konzeptive Übersicht zu neuen Formen der Wissensproduktion und Folgerungen für Programmgestaltung und Evaluierung (Steg, VTI-VTE),
- ein Fallbeispiel für ein Multi-Akteurs-, Multi-Maßnahmen-Programm aus Deutschland, das Programm für Nanotechnologie-Kompetenzzentren in Netzwerkform (Bührer, ISI-Karlsruhe),
- ein weiteres Fallbeispiel aus der Regionalpolitik, nämlich die Bewertung eines Multi-Akteurs-Wettbewerbes in Schleswig-Holstein (Toepel, DIW Berlin),
- ein Beispiel aus Schweden (Materialforschungsprogramm VAMP) gekoppelt mit Systemfragen, wie die "Multiakteure" zusammenwirken (Arnold, Technopolis, UK) sowie
- die Frage nach der politischen Beeinflussbarkeit von Innovationssystemen bzw. Teilsystemen durch Förderprogramme, Institutionen- und Prozessgestaltung (Kuhlmann, ISI-Karlsruhe)

Die Arbeiten werden dazu fortgesetzt und sollen auch in eine Publikation münden. Für Österreich im Allgemeinen und einige Programme (K plus, Regionalpolitik 2000) ergeben sich daraus interessante Ansätze sowohl für das Programmmanagement als auch für Monitoring- und Evaluierungsfragen.

Michael Stampfer, TiG, Grillparzerstrasse 7/8, A-1010 Wien,
Tel. ++43 1/513 26 27
E-mail: michael.stampfer@tig.or.at

PLATTFORM■TECHNOLOGIE■EVALUIERUNG

Die PLATTFORM▪TECHNOLOGIE▪EVALUIERUNG ist eine Initiative des österreichischen BUNDESMINISTERIUMS VERKEHR, INNOVATION UND TECHNOLOGIE. In der Zusammenarbeit von drei österreichischen Forschungseinrichtungen, JOANNEUM RESEARCH, WIFO und FORSCHUNGSZENTRUM SEIBERSDORF und der TECHNOLOGIE IMPULSE GESELLSCHAFT werden Themenstellungen zur Technologieevaluierung erarbeitet und - z.T. mit Einbeziehung namhafter ExpertInnen - in einem Fachkreis diskutiert.

Der Newsletter beinhaltet Fachbeiträge zu Fragen der technologiepolitischen Evaluierung. Die Herausgabe erfolgt in zeitlicher als auch inhaltlicher Abstimmung mit Plattform-Veranstaltungen, um die Synergiewirkungen eines breiten Austauschforums zu nutzen.

Neue Internet-Adresse:

http://www.bmwf.gv.at/4fte/wirtech/techeva/

Für den Inhalt verantwortlich:

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE; A-1010 Wien, Renngasse 5
Dr. Rupert Pichler (rupert.pichler@bmv.gv.at)

TECHNOLOGIE IMPULSE GESELLSCHAFT M.B.H., Grillparzerstr. 7/8, A-1010 Wien
Dr. Michael Stampfer (michael.stampfer@tig.or.at)

JOANNEUM RESEARCH, Institut für Technologie- und Regionalpolitik, A-1040 Wien, Wiedner Hauptstraße 76
Dr. Oliver Fritz (oliver.fritz@joanneum.ac.at)

ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG, A-1103 Wien, PF 91
Mag. Gernot Hutschenreiter (hutsch@wsr.ac.at)

ÖSTERREICHISCHES FORSCHUNGSZENTRUM SEIBERSDORF, Systemforschung Technik-Wirtschaft-Umwelt, A-2444 Seibersdorf
DI Anton Geyer (anton.geyer@arcs.ac.at)

Medienhinweis:
PLATTFORM▪TECHNOLOGIE▪EVALUIERUNG ist ein unregelmäßig erscheinendes offenes Forum zur Diskussion methodischer und inhaltlicher Evaluierungsfragen in der Technologiepolitik. © Wien 2000
Herausgabe und Versand: Joanneum Research, Institut für Technologie- und Regionalpolitik, A-1040 Wien, Wiedner Hauptstraße 76.