
DAGMAR SIMON

Selbststeuerung der Wissenschaft durch Evaluationen?

Optionen und Grenzen institutioneller Forschungsbewertungen

Vor etwas mehr als einer Dekade waren Evaluationen als Bewertungsverfahren im bundesrepublikanischen Wissenschafts- und Forschungssystem nahezu ein Fremdwort – im Gegensatz zu unseren europäischen Nachbarländern wie beispielsweise den Niederlanden oder England, in denen sie seit längerem als ein integraler Bestandteil der Wissenschaftsorganisation verstanden und behandelt werden.¹

Nun droht nach erfolgreicher Implementierung von Bewertungsverfahren im außeruniversitären Forschungsbereich in unterschiedlichen Varianten und mit mehr oder weniger dramatischen Konsequenzen – im Fall der Blauen Liste-Institute wurde und wird letztendlich über die Fortexistenz von Forschungs- und Serviceeinrichtungen entschieden – und der Einführung von Lehr- und auch Forschungsevaluationen auch an den Hochschulen die anfängliche Euphorie, was alles mit diesem Instrumentarium wissenschaftspolitisch bewirkt werden kann, sich in ihr Gegenteil zu verkehren: Von einer zunehmenden „Evaluitis“ ist die Rede, welche die Wissenschaft von ihrer eigentlichen Aufgabe durch die Beanspruchung durch Wissenschaftler/innen als Evaluatoren und die Absorbierung von Forschungseinrichtungen durch die Vorbereitung und Durchführung von Evaluationen abhalte. Die Rede ist von einer Audit-Society und von der Übernahme von Kontrollinstrumenten aus der Finanzwelt, die sich für die Wissenschaft als nicht adäquat erweist.²

Ein schnelles Ende der anfänglichen mit Evaluationen verbundenen Vorstellungen und Hoffnungen? Zumindest werden Anhaltspunkte geboten, um über Ziele und Praktiken von Bewertungsverfahren nachzudenken. Möglicherweise waren und sind die mit Evaluationen verbundenen Erwartungen zu hoch in einer Zeit, in welcher sich der Wissenschafts- und Forschungssektor in einer umfassenden Umbruchssituation befindet, sich mit neuen Herausforderungen für Quali-

1 Vgl. Kieser, A., Going Durch – Was lehren niederländische Erfahrungen mit der Evaluation universitärer Forschung? – In: DBW. 58(1998)2, S. 208 – 224.

2 Charlton, B. / Andras, P., Auditing a Tool of Public Policy: The Misuse of Quality Assurance Techniques in the UK University Expansion. – In: Europea Political Science, 2(2002). S. 24 – 35.

tätskontrolle, Rechenschaftspflicht, Leistungserwartungen und Steuermodelle konfrontiert sieht, die etablierte Praktiken in Frage stellen. Bei diesen in Gang gesetzten Entwicklungen, die unter anderem auch etablierte Machtpositionen und Machtressourcen, insbesondere in den Hochschulen, in Frage stellen und somit besonders „brisant“ sind, wird gern auf Evaluationen als Grundlage für Entscheidungsfindung und -prozesse rekurriert, ob es nun passt oder nicht.

Zudem sind Ziele und Instrumente von Evaluationen an den Hochschulen mit der Einführung neuer Steuerungsinstrumente eng verbunden. U. Schimank geht in diesem Zusammenhang von neuen Governance-Mustern an den Hochschulen aus, die das traditionelle Selbstverwaltungsmodell durch ein Managementmodell ablösen. Dauerhafte Evaluationen mit entsprechenden Konsequenzen für die Ressourcenallokation stellen dabei einen relevanten neuen Steuerungsmechanismus dar.³

Auch die sogenannten vier Säulen der außeruniversitären Forschung – Helmholtz-Gemeinschaft deutscher Forschungseinrichtungen, Fraunhofer-Gesellschaft, Max-Planck-Gesellschaft und die Leibniz-Gemeinschaft – befinden sich im Wandel. Selbstverständliche „Arbeitsteilungen“ beispielsweise zwischen grundlagenorientierter und anwendungsorientierter Forschung werden in Frage gestellt, und der Druck auf die Forschungseinrichtungen, gesellschaftsrelevante Fragen und Problemstellungen in ihre Forschungsprogramme aufzunehmen, wächst seitens der Wissenschaftspolitik.⁴

In den sogenannten Systemevaluationen der großen Wissenschaftsorganisationen wurden Strukturen und Organisationsformen thematisiert, die Instrumente zur Qualitätskontrolle und -förderung kritisch unter die Lupe genommen und vor allem Kooperationsformen zwischen Universitäten und den außeruniversitären Säulen sowie zwischen diesen Forschungseinrichtungen ins Zentrum rücken. Es stehen übergeordnete Gesichtspunkte wie beispielsweise die Einbindung der entsprechenden Wissenschaftsorganisationen in die deutsche Forschungslandschaft – und nicht die Leistungsfähigkeit einzelner Institute – im Vordergrund. So fanden in den letzten Jahren Systemevaluationen auch der grundlagenorientierten Max-Planck-Gesellschaft und der Deutschen Forschungsgemeinschaft sowie der Fraunhofer-Gesellschaft, der Leibniz - Gemeinschaft und jüngst der Helmholtz-Gemeinschaft, also der Großforschungseinrichtungen, statt. Hervorzuheben ist in diesem Zusammenhang, dass zunehmend Fragen der Forschungsorganisation und

3 Schimank, U., Neue Steuerungssysteme an den Hochschulen, Expertise für die Förderinitiative „Science Policy Studies“ des Bundesministeriums für Bildung und Forschung 2002.

4 Vgl. Wissenschaftsrat, Strategische Forschungsförderung. Empfehlungen zur Kommunikation, Kooperation und Wettbewerb im Wissenschaftssystem. Wissenschaftsrat 5654/03, Köln 2003.

Forschungsstruktur als Voraussetzung für erfolgreiche Forschung (und Entwicklung) und damit zusammenhängend auch Qualitätssicherungs- und Qualitätsförderungsverfahren in bezug auf die entsprechenden „Produkte“ – Forschung, wissenschaftliche Dienstleistungen etc. – Aufmerksamkeit gewidmet wird.

Auch in den Thesen des Wissenschaftsrats zur „zukünftigen Entwicklung des Wissenschaftssystems“ (2000) wird akzentuiert, dass die Fähigkeit der Wissenschaftseinrichtungen zur Selbstorganisation erhöht werden muss und Bund und Länder sich aus der Detailsteuerung wissenschaftlicher Einrichtungen zurückziehen sollten. Dagegen werden Instrumente der wissenschaftspolitischen Entscheidungsvorbereitung und -findung wie Evaluation und Prospektion an Bedeutung gewinnen und müssten weiter optimiert werden. Bei der Entwicklung neuer, auch den Bereich der Lehre erfassender Verfahren der Leistungsmessung und Qualitätssicherung sei darauf zu achten, dass diese situationsgerecht und transparent ausgestaltet werden.⁵

Neue beispielsweise veränderte Anforderungen an Wissenschaft und Forschung erhalten insbesondere vor dem Hintergrund der allgegenwärtigen Wissensgesellschaft besondere Brisanz, welche die Ressourcen im Wettbewerb der nationalstaatlichen Ökonomien mobilisieren soll.

Akteure und Interessen

In der bundesdeutschen Diskussion über Evaluierungen und ihre Ziele, Verfahren und Methoden im Wissenschafts- und Forschungssektor lassen sich die unterschiedlichen Akteure leicht ausmachen: Zunächst die Politik, wobei hier ein Trend weg von der Schaffung von Rahmenbedingungen und Inputsteuerungen und hin zur Outputsteuerung zu beobachten ist. Mit Evaluierungen soll eine leistungsorientierte Verteilung von knapper werdender Ressourcen „gewährleistet“ werden; dann die wissenschaftsintermediären Akteure wie beispielsweise der Wissenschaftsrat, die mit ihrem „Gütesiegel“ für die „Qualität“ von wissenschaftlichen Institutionen politische Entscheidungsprozesse über die Zukunft der Forschungseinrichtungen vorbereiten und schließlich die Wissenschaft selbst, die – zumindest in ihrer Mehrheit – von Evaluierungen Qualitätsentwicklungsmaßnahmen zur Verbesserung ihrer Selbststeuerungsfähigkeit erwartet.

Gegen Ende der achtziger Jahre konzentrierten sich in der Bundesrepublik Deutschland die Debatten über Forschungsbewertungen und Evaluierungsmethoden überwiegend auf die Einführung von Evaluationen an den Hochschulen,

5 Wissenschaftsrat, Thesen zur zukünftigen Entwicklung des Wissenschaftssystems in Deutschland. Wissenschaftsrat 4594/00, Köln 2000. S. 7.

die zunächst nur zögernd angegangen wurde. Erst mit der deutschen Einheit, der Begutachtung der außeruniversitären Forschungslandschaft der DDR und der anschließenden Evaluation der Institute der Blauen Liste durch den Wissenschaftsrat gewannen auch die Evaluationen der außeruniversitären, staatlich finanzierten Forschungseinrichtungen an Aufmerksamkeit.

Zwar sind Forschungsbewertungen seit langem ein wichtiger Bestandteil – wenn auch in unterschiedlichem Ausmaß – des wissenschaftlichen Lebens dieser Institute gewesen. Diese Begutachtungen werden meist von den wissenschaftlichen Beiräten der Forschungseinrichtung durchgeführt und haben in der Regel einen hohen Stellenwert für die Selbstvergewisserung und perspektivische Orientierung einer Forschungseinrichtung. Neu ist, dass in jüngster Zeit von allen außeruniversitären, staatlich finanzierten Forschungseinrichtungen der Nachweis von regelmäßigen und systematischen Verfahren der Evaluation erwartet wird, und dass sich alle Einrichtungen in den letzten Jahren externe Evaluationen gefallen lassen mussten.

Seitens der Wissenschaft ist zumindest eine gewisse Skepsis, wenn nicht gar Misstrauen gegenüber der Einführung flächendeckender Evaluationen zu konstatieren. Die Skepsis steht im Kontext mit der Leitidee der vor allem von der Politik geforderten extern gesteuerten Evaluierungen in allen OECD-Ländern, nämlich Leistung transparent zu machen und den Finanzmitteleinsatz effektiver zu gestalten. Signalisiert werden damit zunächst höhere Erwartungen und auch Anforderungen insbesondere politischer Akteure an die Wissenschaft, darüber hinaus deuten diese Forderungen möglicherweise auf einen Vertrauensverlust in die Selbststeuerungsmechanismen der Wissenschaft hin. Die Befürchtung besteht, dass Evaluationen zu politischen Steuerungsinstrumenten werden, welche die Handlungsspielräume der Wissenschaft in ihrer prekären Autonomie berühren.

Aber darüber hinaus herrscht eine gewisse Unsicherheit im Blick auf die unterschiedlichen Zieldimensionen: Sollen sie die Qualität der wissenschaftlichen Leistungen sichern, wobei der Qualitätsbegriff unterschiedliche Vorstellungen von Relevanzkriterien verbirgt oder einen Beitrag zur Legitimation gegenüber staatlichen Zuwendungsgebern leisten, zur Organisationsentwicklung beitragen und das Management der entsprechenden Institution mit Daten und Informationen versorgen oder zur Profilbildung einer Hochschule bzw. außeruniversitären Forschungseinrichtung beitragen oder last but not least für die Adressaten und Nutzer der Forschung Transparenz über die Leistungen herstellen?

In dieser Situation haben besonderes Interesse die Evaluationen der Blauen Liste-Institute durch den Wissenschaftsrat im Zeitraum vom 1995 – 2000 hervorgerufen, zumal diese Bewertungsprozesse zum Teil erhebliche Umstrukturierungen bis zu Schließungen von Einrichtungen zur Folge hatten.

Ein integriertes Evaluationsverfahren

Der Beitrag befasst sich in erster Linie mit den Evaluierungen der Institute der Blauen-Liste.⁶ Diese Forschungs- und Serviceeinrichtungen sind historisch in besonderer Weise durch externe Evaluierungen geprägt, da sich Bund und Land als gemeinsame Träger auf eine externe Bewertung dieser Institute verständigt haben.

Bei den Blauen Liste-Instituten handelt es sich um einen besonderen Typus von Forschungsinstituten, die sich vor allem durch zwei Auffälligkeiten auszeichnen: Sie bewegen sich auf Forschungsfeldern, welche die Bearbeitung sowohl grundlagen- als auch anwendungsorientierter Forschungsfragen – zumeist in multidisziplinären Forschungskontexten – erforderlich machen und darüber hinaus werden die Forschungsaufgaben oftmals um Beratungs- und Serviceleistungen für wissenschaftliche und außerwissenschaftliche Bezugsgruppen ergänzt. Dieser Typus hat seit einigen Jahren in der wissenschaftspolitischen und wissenschaftssoziologischen Diskussion eine besondere Aufmerksamkeit erlangt: Durch den Bedeutungsgewinn von wissenschaftlich-technischem Wissen im internationalen ökonomischen Wettbewerb ist das Verhältnis von Technik und Wissenschaft (vor allem auf den Gebieten der sogenannten Schlüsseltechnologien) und darüber hinaus das Verhältnis von grundlagen- und anwendungsorientierten Forschungen zu einem relevanten Thema geworden.

Ins Zentrum des wissenschaftspolitischen Diskurses gerät zunehmend die Frage nach den optimalen Voraussetzungen („Produktionsformen“) für innovative Forschung. Organisationsformen etablierter Forschungseinrichtungen finden in Evaluationen wachsende Aufmerksamkeit. Insbesondere mit Blick auf die zukünftige öffentliche Forschungsförderung wird die steigende Bedeutung außerwissenschaftlicher Zusammenhänge und sozialökonomischer Prioritäten akzentuiert sowie die Notwendigkeit betont, an gesellschaftlichen Problemlagen orientierte Forschungsfragen unter Einbeziehung von Anwendern zu bearbeiten.⁷

Daran anknüpfend wird im Rahmen der Wissenschaftsforschung darüber gestritten, ob wir es tatsächlich mit einer veränderten Form der Wissensproduktion zu tun haben, wie es Gibbons und andere in ihrer Studie „New Production of Knowledge“ behaupten und sie mit Modus 2 beschreiben.⁸ Demnach war ursprünglich ein Referenzmodell dominant, bei dem Wissenschaftler in einem primär selbstbezüglichen Diskurs Deutungen und Empirieprogramme entwickelt

6 Röbbcke, M. / Simon, D., Reflexive Evaluation. Ziele, Verfahren und Instrumente der Bewertung von Forschungsinstituten. Berlin: edition sigma 2001.

7 Vgl. Wissenschaftsrat, Strategische Forschungsförderung. Empfehlungen zur Kommunikation, Kooperation und Wettbewerb im Wissenschaftssystem. Wissenschaftsrat 5654/03, Köln 2003.

und verabschiedet haben. Die Gesellschaft delegiert somit keine Fragen an die wissenschaftliche Gemeinde, sondern diese definiert die Problemstellungen und produziert Ergebnisse, ohne die Verantwortung für eine adäquate Perzeption zu übernehmen. Dieser Typus soll weitgehend durch ein neues „Produktionsmodell“ abgelöst worden sein. Forscher arbeiten mit politischen und wissenschaftlichen Auftraggebern und/oder Anwendern zusammen und erarbeiten die Problemlagen in einer Gemengelage, bei der die wissenschaftlichen Disziplinen keine dominierende Referenz mehr darstellen. Die Thesen bedürfen noch der empirischen Belege; sie sind hier auch als weiterer Hinweis zu verstehen, sich mit den „Produktionsformen“ wissenschaftlicher Arbeit näher zu befassen.

Befasst man sich mit den Zielen, Verfahren und Kriterien von institutionellen Evaluationen von Forschungsinstituten, wird deutlich, dass erstens Evaluationen grundsätzlich ein breites Spektrum von Zielen haben können, das sich von forschungspolitischen Interventionen über die Generierung von Informationen für die Ressourcenallokation bis zu internen Prozessen der Qualitätsförderung erstreckt. Im Fall der Blauen Liste-Institute sind die Ziele der Evaluation erheblich durch wissenschaftspolitische Vorgaben beeinflusst, denn – wie gesagt – nach den Vorstellungen von Bund und Ländern sollen im Rahmen von regelmäßigen Evaluationen wissenschaftspolitische Prioritäten gesetzt werden, indem über die Weiterförderung der einzelnen Institute entschieden wird. Zugleich sollen die Evaluationen einer kontinuierlichen und systematischen Qualitätssicherung dienen – ein Ziel, das aus der Perspektive der Institute nachdrücklich unterstützt wird. Möglicherweise entstehen dabei aber Zielkonflikte, denn der hohe Leistungsdruck einer Evaluierung, deren Ziel die „Flexibilisierung“ der Blauen Liste ist, lässt sich nur schwer mit einer Qualitätssicherung und -förderung vereinbaren, die auf einer selbstkritischen Stärke-Schwäche-Analyse basiert.

Zweitens ging es um die Frage nach den geeigneten Verfahren zukünftiger Evaluationen. Es gehört zu den Grunderkenntnissen der Evaluationsforschung, dass die unterschiedlichen Ziele einer Evaluation auch an unterschiedliche Verfahren gebunden werden müssen. Außerdem haben die Wahl und Ausgestaltung der Verfahren einen entscheidenden Einfluss darauf, ob es den Forschungseinrichtungen gelingt, Evaluationen als einen Prozess der Selbstkontrolle zu gestalten und eine Überwältigung durch wissenschaftsfremde Interventionen zu verhindern.

Drittens geht es um die Bewertungskriterien, nach denen die Blaue Liste-Einrichtungen beurteilt werden. Während sich die Suche nach geeigneten Verfahren

8 Gibbons, M. / Limoges, C. / Nowotny, H. / Schwartzman, S. / Scott, P. / Trow, M., *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies.* London-New Delhi 1994.

noch auf Modelle stützen kann, die bereits in anderen Kontexten erfolgreich entwickelt worden sind, gestaltet sich die Ermittlung von angemessenen Indikatoren schwieriger. Dafür gibt es mehrere Gründe: Erstens sind die Verlässlichkeit und Gültigkeit von Forschungsindikatoren umstritten, zweitens sind mit der Notwendigkeit, die erhobenen Daten sachkundig und unabhängig zu interpretieren, neue Verfahrensprobleme verbunden, und drittens müssen überzeugende Indikatorenprogramme für sehr unterschiedliche Einrichtungen konstruiert werden. Denn die Aufgabenprofile der Blauen Liste-Institute sind durch verschiedene Dimensionen gekennzeichnet: durch unterschiedliche Gegenstandsbereiche, Forschungstypen und Anwendungskontexte. Die Einrichtungen orientieren sich zum einen an innerwissenschaftlich generierten relevanten Kriterien und der Anerkennung durch die scientific community; sie sind also auf das spezifische Belohnungssystem der Wissenschaft und den Reputationserwerb angewiesen. Zum anderen müssen die Institute wissenschaftspolitischen Erwartungen gerecht werden, wozu in jüngerer Zeit auch eine verstärkt geforderte Ausrichtung der Forschung auf marktfähige Produkte, Verfahren und Dienstleistungen gehört. Für eine Evaluation dieser Institute sind daher differenzierte Bewertungs- bzw. Fragenkataloge erforderlich.

Zur Sicherung, Kontrolle und Weiterentwicklung der Forschungs- und Dienstleistungen wird ein integriertes Modell von externer Evaluation und Selbstevaluation vorgeschlagen, das auf eine Qualitätsförderung der Institutsleistungen abzielt und nicht ausschließlich auf Maßnahmen zur Qualitätskontrolle und -sicherung.

Nun ist die Frage nach einer adäquaten Definition der Qualität von Forschung nicht leicht zu beantworten. Damit zusammen hängt auch die Frage nach den Akteuren, die Qualität definieren und redefinieren sowie den Instrumenten, die Qualität fördern. Ein amerikanischer Evaluationsforscher stellt so auch einmal schon etwas entnervt die Frage: „What the hell is quality?“

Van Vught hebt beispielsweise den „multidimensionalen“ und „subjektiven“ Charakter des Qualitätsbegriffs hervor, der zeitweilig den Charakter eines politischen Konzeptes aufweist, das in den letzten Jahren von Bildungspolitikern verstärkt be- und genutzt wird.⁹ Jedoch impliziert der Qualitätsbegriff, dass nicht nur nach den Stärken der Lehre und Forschung – in diesem Fall – gefragt wird, sondern dass gerade auch Schwächen aufgedeckt werden und der Umgang mit ihnen thematisiert wird. Die Deutung des Qualitätsbegriffs ist je nach Akteuren und spezifischen Interessenkonstellationen unterschiedlich, insbesondere hin-

9 van Vught, F., *The Humboldtian University under Pressure. New Forms of Quality Review in Western European Higher Education.* – In: *Inside Academia. New Challenges for the Academic Profession.* Ed. by P. Maassen and F. van Vught. Utrecht 1994. S. 185 ff.

sichtlich der mit Evaluationen verbundenen Zielvorstellungen und ihrer Konsequenzen – „quality as a moving target“.¹⁰

In der Diskussion über Evaluationsziele im Wissenschafts- und Forschungssektor und insbesondere der Hochschulen stehen die Ziele „accountability“ und „improvement“ im Zentrum; ihr Verhältnis zueinander wird unterschiedlich charakterisiert. In der niederländischen Evaluationsforschung werden „accountability“ und „improvement“ als entgegengesetzte Pole – Scylla und Charybis – gesehen¹¹, von anderer Seite wird den meisten nationalen Evaluationssystemen eine Verbindung von öffentlicher Rechenschaftslegung und einer an den Zielen und Konzepten der Hochschule orientierten Qualitätsentwicklung, wenn auch unterschiedlicher Ausprägung, zugeschrieben.¹²

„Quality improvement“ wird im internationalen Vergleich in der Diskussion um Instrumente und Methoden zur Steigerung der Leistungsfähigkeit der Hochschulen als Weiterentwicklung des „quality assessment“ betrachtet und favorisiert. An den Hochschulen wird darüber hinaus in der Diskussion über Qualitätskontrolle und -förderung auf vielfältige Instrumente wie die Entwicklung langfristiger Strategien, Organisations- und Managemententwicklung und die Reform der Leitungsstrukturen rekurriert. Ihnen liegt der Gedanke zugrunde, dass Qualität eine kontinuierliche Verbesserung erfordert, die auch in anderen Qualitätskonzepten, wie Zertifizierungen, zum Ausdruck kommt.

Für außeruniversitäre Forschungsinstitute liegt eine besondere Chance in solchen Evaluationen, die nicht nur den Status quo bewerten, sondern die Weiterentwicklung der Einrichtung und ihrer Forschungsaufgaben als ihre Aufgabe ansehen. Keineswegs ausgeschlossen sind damit Entscheidungen, die auf die Einstellung von Forschungsbereichen, -programmen oder anderer Institutsleistungen oder sogar der Einrichtung insgesamt hinauslaufen. In einer Folgeevaluation ist dann zu überprüfen, ob und in welcher Weise das Institut die Empfehlungen umgesetzt hat. Damit kann auch die Lernfähigkeit des Instituts als eine entscheidende Voraussetzung zur Steigerung der Leistungsfähigkeit gefördert werden.

Der Gedanke einer kontinuierlichen Verbesserung der Qualität der Forschung wird am ehesten in einem integrierten Modell von externer Evaluation und Selbstevaluation realisiert. Die „Qualitätsdebatte“ macht zweierlei deutlich: Die

- 10 Felt, U., Evaluationen im wissenschaftspolitischen Kontext. – In: Qualitätsförderung durch Evaluationen? Ziele, Aufgaben und Verfahren von Forschungsbewertungen im Wandel. Hrsg. v. M. Röbbecke und D. Simon. WZB Discussion Paper P99 – 003. Berlin 1999. FS. 11 ff.
- 11 Vroeijsenstijn, A.I., Improvement and Accountability: Navigating between Scylla and Charybis. Guide for External Quality Assessment. – In: Higher Education. London-Bristol 1995.
- 12 Barz, A. / Carstensen, D. / Reissert, R., Lehr- und Evaluationsberichte als Instrumente zur Qualitätssicherung, Bestandaufnahme der aktuellen Praxis. Gütersloh 1997.

unterschiedlichen Ziele „accountability“ und „improvement“ werden externen Evaluierungen einerseits und Selbstevaluationen andererseits zugeordnet.¹³ Darüber hinaus werden Fragen der Qualitätssicherung und Qualitätsförderung im direkten Zusammenhang mit der Existenz und dem Ausbau von Selbstkontroll- und Selbstevaluationssystemen aufgeworfen und diese als unabdingbare Voraussetzung für den effektiven Einsatz von qualitätssichernden Maßnahmen analysiert.¹⁴

Externe Evaluationen sollten sich sowohl auf die Bewertung der Forschungs- und Dienstleistungen einer Forschungseinrichtung beziehen, gemessen an seinem Aufgaben- und Leistungsprofil, als auch auf die Überprüfung der Existenz und Ausgestaltung von institutsinternen Maßnahmen zur Qualitätskontrolle und -förderung, also der Bedeutung und Qualität von Verfahren der Selbstevaluation.

Unter den spezifischen wissenschaftspolitischen Rahmenbedingungen – die externe Evaluation ist gekoppelt an eine wissenschaftspolitische Entscheidung über die Weiterförderung des Instituts – werden durch externe Evaluationen drei Aufgaben wahrgenommen: die der Qualitätskontrolle, Qualitätssicherung und Qualitätsförderung.

Die Qualitätskontrolle bezieht sich auf zwei Referenzsysteme: Die Leistungen der jeweiligen Einrichtung müssen vor dem Hintergrund des Aufgaben- und Leistungsprofils sowie der selbstgesteckten Ziele beurteilt werden. Darüber hinaus ist danach zu fragen, ob sie (die Ziele) den allgemeinen (wissenschaftspolitischen) Anforderungen an diese außeruniversitären Forschungseinrichtungen (zum Beispiel Kooperationen mit den Hochschulen, Nachwuchsförderung etc.) entsprechen.

Bei der Qualitätssicherung steht die Überprüfung der Instrumente und Verfahren interner Kontrollen – ihrer Funktionalität und Reichweite – im Hinblick auf eine externe Validierung ihrer Qualität im Vordergrund. Sie sind ein entscheidender Bestandteil zur Beurteilung der Institutsperformanz. Mit der Qualitätsförderung wird versucht, aus einer Außenperspektive Fehlentwicklungen zu identifizieren, Lösungswege zu eruieren und Empfehlungen zur Weiterförderung des Instituts zu geben.

- 13 Westerheijden, D.F. / Brennan, J. / Maassen, P.A.M., *Changing Contexts of Quality Assessment. Recent Trends in West European Higher Education*. Utrecht 1994; van Vught, E., *The Humboldtian University under Pressure. New Forms of Quality Review in Western European Higher Education*. – In: *Inside Academia. New Challenges for the Academic Profession*, a.a.O.; Hrsg. v. Altrichter, H., / Schratz, M., *Qualität von Universitäten. Evaluation: Impulse für Innovation?* Innsbruck 1992.
- 14 Zum Beispiel Kells, H.R., *Self-Regulation in Higher Education. A Multi-National Perspective on Collaborative Systems of Quality Assurance and Control*. London 1992.

Interne Evaluationen sollen dazu beitragen, den Leistungsstand der jeweiligen Einrichtung zu erfassen, Stärken und Schwächen zu analysieren und Perspektiven der zukünftigen Entwicklung zu entfalten. Die Institutionalisierung eines solchen Verfahrens wird primär als kollektiver Lernprozess und weniger als eine Kontrollroutine verstanden. Das Verfahren ist als ein regelmäßiger, nicht auf Einmaligkeit angelegter, institutsinterner Verständigungs- und Entscheidungsprozess zu konzipieren.

Im Sinne der Förderung der Lernfähigkeit von Institutionen mit dem Ziel der Verbesserung und Optimierung ihrer „Produkte“ sowie ihrer strukturellen und organisatorischen Voraussetzungen können Evaluationen dazu beitragen, der allgegenwärtigen Gefahr der Routinisierung ihrer Aufgaben und der institutionellen „Erstarrung“ durch einen umsetzungsorientierten Diskussionsprozess zu begegnen und so die Selbststeuerungsfähigkeit erhöhen. Voraussetzung ist die aktive Einbeziehung des Institutspersonals in diesen Prozess sowie die Herstellung eines gemeinsamen Grundverständnisses über Ziele, Aufgaben und Instrumente.

An dieser Stelle kann nicht umfassend auf die Diskussion über qualitative und quantitative Kriterien eingegangen werden. Nur soviel dazu: Üblich ist es seit langem in Wissenschaft und Forschung, dass ihre „Produkte“ (zum Beispiel Publikationen) auf der Grundlage von qualitativen Kriterien, vor allem durch peer review, beurteilt werden. Erst in jüngerer Zeit spielen quantitative Kriterien wie die erfolgreiche Einwerbung von Drittmitteln oder die Publikationshäufigkeit in den Evaluationen eine relevante Rolle. Bei den untersuchten Forschungseinrichtungen ergibt sich insbesondere die Problematik, dass aufgrund ihrer unterschiedlichen Aufgaben und Ziele aufgrund vielfältiger und unterschiedlicher Bezugsgruppen ihre Leistungen („Produkte“) nicht ausschließlich nach akademischen Kriterien (zum Beispiel Drittmittelinwerbung bei der Deutschen Forschungsgemeinschaft, Publikationen in sog. refereed journals etc.) beurteilt werden können. Es stellt sich vielmehr die Aufgabe, zum Beispiel Dienstleistungen wie Beratung politischer Akteure, die Bereitstellung von technologischen Produkten für die Industrie adäquat zu beurteilen (zum Beispiel durch die Unterscheidung zwischen Dienstleistungsprozessen und -produkten, Qualitätsermittlung durch Kundenbefragungen etc.).

Ein Deskriptorenprogramm, in dem Bewertungskriterien zur Geltung kommen, muss so konzipiert sein, dass die jeweils spezifischen Leistungen und Aufgaben der verschiedenen Einrichtungen erfasst und begutachtet werden. Dies reicht jedoch nicht aus; vielmehr ist zugleich zu klären, welche Relevanz den begutachteten Tätigkeiten und Leistungen im gesamten Aufgabenspektrum der jeweiligen Einrichtung zukommt. Beispielsweise wird in den derzeitigen Evaluationen den Dienstleistungen ein weit geringerer Stellenwert beigemessen als den Forschungsleistungen. Es geht also um die Gewichtung und Interpretation einzelner Des-

kriptoren. Dabei sind vier Ebenen zu unterscheiden: die Gewichtung und Interpretation eines einzelnen Deskriptors in verschiedenen Einrichtungen, die Gewichtung der Dienstleistungen im Verhältnis zu den Forschungsaufgaben, die Gewichtung unterschiedlicher Deskriptoren innerhalb eines Instituts sowie ihre Gewichtung und Interpretation über eine längere Zeitspanne.

Bei den Deskriptoren wird zwischen Input (Gesamtheit von äußeren Einflüssen und internen Voraussetzungen für den Prozess der Leistungserstellung), Output (Ergebnisse der Leistungserstellung) und Troughput (Prozesse und Strukturen, die der Umwandlung des Input in den angestrebten Output dienen) unterschieden. Die Bedeutung der letztgenannten Kategorie wird durch die neuere Diskussion in der Wissenschaftspolitik über Strukturen, Organisations- und vor allem Kooperationsformen in der Forschung nachdrücklich unterstrichen.

Die Leistungsfähigkeit eines Forschungsinstituts ist im hohen Maße von den organisatorischen Rahmenbedingungen der Forschung beeinflusst. Es handelt sich um betriebsförmige Organisationen, das heißt Struktur- und Steuerungsentscheidungen sind von erheblicher Bedeutung für die wissenschaftliche Produktivität. Fragen der Personalplanung und der Stellenbesetzung sowie die Bedeutung und Funktionalität von internen Beratungs- und Entscheidungsgremien spielen dabei eine besondere Rolle in den Blauen Liste-Instituten.

Jede Forschungseinrichtung sieht sich mit der Aufgabe konfrontiert, strukturelle oder verfahrensmäßige Lösungen zur Sicherung und Förderung der eigenen Leistungsfähigkeit zu entwickeln. Dazu gehören die regelmäßige und kontinuierliche Kontrolle der Qualität des Outputs und der selbstkritische Umgang mit den eigenen Struktur-, Organisations- und Steuerungsentscheidungen. Die Qualitätskontrolle ist einer der wichtigsten Bestandteile des „Throughput“ und zugleich das Scharnier zwischen interner und externer Evaluation. Im besten Fall handelt es sich bei der externen Evaluation um eine Metaevaluation der internen Evaluation.

Die Evaluation der Blauen Liste-Institute machte deutlich, dass es kein Organisationsmodell gibt, das für alle Forschungseinrichtungen optimal ist, wohl aber sollten die Organisationsstrukturen den Besonderheiten der unterschiedlichen Einrichtungen adäquat sein. Die Ermittlung von jeweils angemessenen Organisations- und Steuerungsstrukturen kann sich jedoch derzeit erst auf einige wenige organisations- oder wissenschaftssoziologische Untersuchungen bzw. Beiträge aus der Managementforschung zu den Voraussetzungen erfolgreicher Forschung stützen¹⁵; hier besteht weiterhin Nachholbedarf, um die in der aktuellen wissenschaftspolitischen Debatte herausgehobene Bedeutung der „spezifischen Arbeitsweisen, der strategischen Handlungsfähigkeit und den Verfahren zur Qualitätssicherung“ im Wissenschafts- und Forschungssektor eine empirische Fundierung zu geben.¹⁶

Die Forschungslandschaft ist insgesamt mit neuen An- und Herausforderungen auf verschiedenen Ebenen konfrontiert: mit der Generierung neuer Forschungsthemen, die dem gesellschaftlichen und insbesondere wirtschaftlichen Innovationsbedarf nachkommt, mit strukturellen Veränderungen, die beispielsweise neue – nationale und internationale – Kooperationsformen und Vernetzungen, insbesondere zwischen Hochschulen und außeruniversitären Einrichtungen, hervorbringen kann, mit veränderten Ansprüchen der Gesellschaft bezüglich öffentlicher Rechenschaftslegung über die Verwendung von Forschungsmitteln sowie mit neuen Formen der Qualitätssicherung durch die Einführung interner Kontroll- und Steuerungsmechanismen. Die Forschungsorganisationen sind gefordert, ihre Selbstorganisation zu professionalisieren, um sich in dem institutionellen Prozess der Verflechtung unterschiedlicher Interessen und Erwartungen behaupten zu können.

Versucht wurde, die Konturen eines Evaluationsverfahrens zu skizzieren, das auf die Lernfähigkeit von Institutionen setzt mit dem Ziel der Qualitätsförderung als kontinuierlicher Aufgabe von außeruniversitären Forschungsinstitutionen und somit zu einer wissenschaftsbasierten und wissenschaftsadäquaten Selbststeuerung beitragen kann. Dieses Modell beruht auf einem Zusammenspiel von externen und internen Evaluationen.

Ein Akzent wurde insbesondere auf die Selbstevaluationen der Institute gesetzt, um in dem komplizierten Geflecht der Interessen externer Akteure sie in ihrer Fähigkeit zu stärken, als eigenständiger Akteur aufzutreten. Dazu bedarf es eines selbstkritischen Umgangs und entscheidungsorientierter Reflexions- und Kontrollprozesse, die nicht als Ausnahmezustand, sondern als integralen Bestandteil des Forschungsprozesses verstanden werden sollten. Es kann deshalb auch ein primäres Ziel von externen Evaluationen sein, Selbstevaluationen zu stimulieren und zu qualifizieren.

Mittlerweile sind in der Evaluationsforschung und -praxis Bemühungen im Gange, Standards der Selbstevaluation festzulegen – u.a. vor dem Hintergrund, dass zukünftig nur diejenigen Organisationen und Unternehmen antragsberechtigt für öffentlich finanzierte Maßnahmen sind, deren System der Qualitätssicherung durch eine Zertifizierungs- oder Akkreditierungsstelle anerkannt ist.¹⁷

- 15 Zum Beispiel: Hasse, R., *Organisierte Forschung. Arbeitsteilung, Wettbewerb und Networking in Wissenschaft und Technik*. Berlin: edition sigma 1996; von Stebut, N., *Eine Frage der Zeit? Zur Integration von Frauen in die Wissenschaft. Eine empirische Untersuchung der Max-Planck-Gesellschaft*. Opladen: Leske + Budrich 2003; Wimbauer, Ch., *Organisation, Geschlecht, Karriere. Fallstudien aus einem Forschungsinstitut*. Opladen: Leske + Budrich 1999.
- 16 Vgl. Bericht der Internationalen Kommission zur Systemevaluation der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft, *Forschungsförderung in Deutschland*. Hannover 1999.

Es ist möglicherweise an der Zeit, die bislang gesammelten Erfahrungen von Evaluierungen in Wissenschaft und Forschung hinsichtlich ihrer intendierten und vor allem nicht intendierten Folgen einer „Metaevaluation“ zu unterziehen. Alfred Kieser hat schon vor einigen Jahren aus der Evaluation der Betriebswirtschaftslehre an den niederländischen Hochschulen folgende Lehren für zukünftige Verfahren gezogen: Evaluationen der Forschung müssen die Besonderheit der Forschungsprogramme und -projekte und die Anwender im Evaluationsprozess berücksichtigen; ein Mix von Evaluationsmethoden ist ausschließlich quantitativen Wertungen vorzuziehen; interne Evaluationen haben Vorrang vor externen; Transparenz muss für alle Beteiligten hergestellt und die Evaluationen als ein Prozess gestaltet werden, in dem Evaluierete und Evaluierer miteinander lernen.¹⁸ Der Wissenschaftsrat hat in dem Reflexionsprozess seiner Praktiken ähnliche Punkte akzentuiert.¹⁹

Vor allem geht es auch um die Angemessenheit des jeweiligen Verfahrens für die Ziele der Qualitätskontrolle und -förderung; um nicht konstatieren zu müssen: „It was a methodological error to use audit technologies to award grades to British universities. The fault presumably arose from the initial dishonesty of announcing that quality assurance would be used to raise standards, which implies a quantitative system of grading. A proper quality assurance system would maintain minimum consistent standards, but it is not of itself a system for continually cranking-up standards”.²⁰

Soll eine Evaluation weder eine *l'art pour l'art* – Veranstaltung noch eine bürokratische Kontrollroutine werden, sondern ein für die entsprechende Institution produktiver Prozess, der die Chancen auf Selbststeuerung erhöht, können solche Verfahren als eine Interaktion verstanden werden, zu dem Evaluateure und Evaluierete jeweils beitragen. Dementsprechend müssen Evaluierungen klare mit dem Bewertungsprozess verbundene Zielvorstellungen vermitteln und für alle Beteiligten transparent und partizipatorisch gestaltet werden. Somit kann auch die Akzeptanz von Evaluationen bei allen Beteiligten erhöht und damit die Chance, dass intendierte Veränderungsprozesse tatsächlich eingeleitet werden.

17 Müller-Kohlenberg, H. / Beywl, W., Standards der Selbstevaluation – Begründung und aktueller Diskussionsstand. – In: Zeitschrift für Evaluation. 1(2003), S. 67.

18 Kieser, A., Going Dutch – Was lehren niederländische Erfahrungen mit der Evaluation universitärer Forschung? – In: DBW. 58(1998)2, S. 222.

19 Wissenschaftsrat, Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates. Wissenschaftsrat 5375/02, Köln 2002.

20 Charlton, B. / Andras, P., Auditing a tool of Public Policy: The Misuse of Quality Assurance Techniques in the UK University Expansion. – In: Europea Political Science. 2(2002), S. 32.

Gesellschaft für
Wissenschaftsforschung



Klaus Fischer
Heinrich Parthey (Hrsg.)

**Evaluation
wissenschaftlicher
Institutionen**

Wissenschaftsforschung
Jahrbuch 2003

Sonderdruck

Mit Beiträgen von:

Wolfgang Biedermann • Manfred Bonitz

Klaus Fischer • Siegfried Greif

Frank Havemann • Marina Hennig

Heinrich Parthey • Dagmar Simon

Roland Wagner-Döbler

Wissenschaftsforschung
Jahrbuch **2003**

Deutsche Nationalbibliothek
Evaluation wissenschaftlicher Institutionen
: Wissenschaftsforschung Jahrbuch 2003 /
Klaus Fischer; Heinrich Parthey (Hrsg.). -
Berlin: Gesellschaft für Wissenschaftsforschung 2011.
ISBN: 978-3-934682-57-3

2. Auflage 2011
Gesellschaft für Wissenschaftsforschung
c/o Institut für Bibliotheks- und
Informationswissenschaftswissenschaft
der Humboldt-Universität zu Berlin
Unter den Linden 6, D-10099 Berlin
<http://www.wissenschaftsforschung.de>
Redaktionsschluss: 15. März 2011
This is an Open Access e-book licensed under
the Creative Commons Licence BY
<http://creativecommons.org/licenses/by/2.0/>