

SCHWERPUNKTTHEMA

Evaluation von Forschung

Einführung in den Schwerpunkt

Das Motiv, sich dem Thema „Evaluation von Forschung“ in einem Schwerpunkt zu widmen, ist u. a. ein relativ eigennütziges: Mit dem Aufbau entsprechender Strukturen in der HGF (der Helmholtz-Gemeinschaft Deutscher Forschungszentren, die die frühere Arbeitsgemeinschaft der Großforschungseinrichtungen abgelöst hat) und dem Übergang auf eine programmorientierte Förderung waren auch geeignete Ansätze zu entwickeln, wie die zentrenübergreifend eingerichteten Programme evaluiert und in ihrer weiteren Umsetzung verfolgt und gesteuert werden können („Controlling“). Natürlich wurden die Forschungsprogramme der Zentren auch schon vorher evaluiert, aber die neue Runde erforderte einen eigenen Ansatz und Ausschau nach vorhandener Erfahrung. Wenn eine ganze Sparte des deutschen Wissenschaftssystems nach einem neuen Ansatz evaluiert wird, hat dies auch Bedeutung über den Bereich der HGF hinaus. Über umfangreiche Evaluationserfahrungen verfügt selbstverständlich der Wissenschaftsrat (WR)¹. Er hat in der Vergangenheit nicht nur einzelne Institute evaluiert, sondern ganze Wissenschaftssparten in Form sog. „Systemevaluationen“ – so für die Institute der Blauen Liste, die sich als Wissenschaftsgemeinschaft Leibniz (WGL) eine engere Kooperationsstruktur gegeben haben (WR 2001a), oder auch für die HGF (WR 2001b). Schließlich hat der Wissenschaftsrat nicht nur bei den Instituten der Blauen Liste solche mit einem politikberatenden Auftrag evaluiert, sondern im letzten Jahr auch einige Landesinstitute (zuletzt die Akademie für Technikfolgenabschätzung in Stuttgart), so dass es auch für ITAS als einem solchen politikberatenden Institut von Interesse ist zu sehen, wie eine solche Evaluation angelegt wird.

Beim gewählten Thema hat das Schwerpunktheft dieses Mal einen etwas anderen Auf-

bau und Zuschnitt, weil auch Erfahrungen mit Evaluationen aufgenommen werden sollten. Nehmen wir die vorliegende Einführung hinzu, dann gruppieren sich die Beiträge des Heftes wie folgt: 1-3-2-3-1; nach der Einführung folgen drei Beiträge, die Überblickscharakter haben, zunächst zum übergeordneten technik- und forschungspolitischen Kontext, dann zu bibliometrischen Ansätzen und mit TA-SWISS das praktische Beispiel der Evaluation einer TA-Einrichtung.

In der nächsten Gruppe finden sich zwei Beiträge aus dem Bereich der Evaluationen des Wissenschaftsrates, zunächst der Evaluationsansatz des WR, wie er über die Jahre entwickelt und immer wieder revidiert wurde, und einen Beitrag aus einem betroffenen Institut.

Die dann folgenden drei Beiträge sind allesamt Erfahrungsberichte aus den jüngsten Evaluationen in der HGF, womit eine Nähe zu forschungspolitisch hochbrisanten Begutachtungen gegeben ist, so dass eine distanzierte Reflexion sich nicht einfach gestaltet.

Am Ende steht ein Evaluationsansatz aus einem neu eingerichteten Forschungsfeld der sozial-ökologischen Forschung, der auf neue Weise versucht zusammenzubringen, was Kuhlmann in seinem Überblick als Themen anschlägt: „Leistungsmessung oder Lernmedium?“

Da in den einzelnen Beiträgen mit immer wieder neuen Akzenten ausgeführt wird, was unter Evaluation verstanden werden kann und wie man sie machen kann, muss sich diese Einleitung mit solchen Abgrenzungsfragen nicht belasten. Nur auf ein Kernproblem sei hier hingewiesen, das u. E. verdeutlicht, warum Evaluationen ein so schwieriges und delikates Geschäft sind. Evaluationen sind keine einfachen Prüfvorgänge (in diesem Punkt widersprechen wir Stockmann 2000, S. 11) oder voll operationalisierte Messungen; Evaluationen sind komplexe Beurteilungsverfahren, die auf komplexe Sachverhalte (Forschungsprogramme, Institutionen, ganze Innovationssysteme) angewandt werden, und dies unter ganz bestimmten Zielsetzungen und Kriterien. Bei einer solchen Betonung springt sofort die pragmatische Dimension ins Auge, d. h. das komplexe Instrumentarium wird gehandhabt, die Ziele werden ausgehandelt, das Ergebnis steht nicht von selbst da, sondern muss interpretiert, ausgelegt und dann

wieder angewandt, umgesetzt werden. Dabei kann man sich zwar auf ein bestimmtes Vorgehen einigen, die Ziele absprechen, sich in der Interpretation der Ergebnisse abgleichen, d.h. einen Konsens finden, wie eine Sache anzupacken ist; aber es ist klar, dass man sich über jedes Detail einer Evaluation auch streiten kann. Evaluation ist im Kern auch ein „politischer“ Sachverhalt. Evaluationen, da stimmen wir Stockmann (200, S. 16) wieder zu, haben Gestaltungs-, Kontroll-, Steuerungs- und Bewertungsfunktionen, und Evaluationsforschung bewegt sich „in einem Minenfeld politischer, administrativer und gesellschaftlicher Interessen“, wie er mit einem Zitat von Hellstern und Wollmann (1980, S. 61) festhält, impliziert also eine Dualität, „die sich darin ausdrückt, dass sie einerseits Teil der empirischen Sozialforschung [ist] ..., aber andererseits auch Teil des politischen Prozesses ..., den sie selbst mit ihren Ergebnissen beeinflusst und umgekehrt als Instrument zur Entscheidungsfindung für die politische Steuerung wissenschaftsfremden Anforderungen ausgesetzt ist ...“ (S. 17). Von solchen politischen Implikationen ist selbst das „peer reviewing“ nicht frei, wie Lawrence (2003) aufgrund eigener umfangreicher Erfahrungen als Gutachter und Editor betont und er deshalb zu Recht von „The politics of publication“ spricht. Wir gehen nun die einzelnen Beiträge durch, geben eine kurze Charakterisierung und heben den einen oder anderen Gesichtspunkt hervor.

Der einführende Beitrag von **Stefan Kuhlmann** hat die Funktion des Überblicks: **Evaluation in der Forschungs- und Innovationspolitik** und er stellt das Thema unter zwei Pole: **Leistungsmessung oder Lernmedium?** Kuhlmann zeichnet die Entwicklungslinien nach, charakterisiert dabei die deutsche Forschungslandschaft und ordnet die unterschiedlichen wissenschaftsinternen und wissenschaftsexternen Evaluationsansätze nach einem Drei-Schalenmodell: den Kern bilden die individuellen Forschungsleistungen, die mit Hilfe von wissenschaftsinternen Instrumenten, vor allem dem „peer reviewing“, evaluiert werden; die nächste Schale (und auch Ebene) sind forschungs- und innovationspolitische Programme, z. B. seitens der staatlichen Forschungspolitik, mit denen bestimmte Ziele verfolgt und Entwicklungen beeinflusst werden sollen; hier kommen dann wirkungsorientierte Evaluations-

ansätze von meist externen Evaluatoren zum Einsatz; schließlich geht es bei der dritten Schale um Institutionen und auch ganze Wissenschaftssparten (wie Max Planck Institute, Fraunhofer Institute oder eben HGF); solche „Systemevaluationen“ des Wissenschaftsrates wurden einleitend schon erwähnt.

Kuhlmann schließt eine kurze Diskussion zur Frage der „Objektivität und Reichweite von Evaluationsverfahren“ mit der Warnung: „Vorsicht und Sorgfalt sind angebracht.“ Die Zusammenhänge zwischen gefördertem Programm und den Erträgen (in vielfacher Hinsicht) sind „komplex und [können] keinesfalls als schlichtes Input/Output-Modell konstruiert werden.“

Schon sehr früh sind in der Evaluationsforschung zwei Formen bzw. zwei funktionale Pole unterschieden worden, nämlich auf der einen Seite eine „summative“ Evaluation, die die Effekte und Wirkungen bilanziert und die letztlich als eine Art von Leistungsmessung zu bewerten ist, und auf der anderen Seite die „formative“ Evaluation, bei der die Analyse- und Bewertungsanstrengungen darauf gerichtet sind, die Entwicklung (z. B. eines Reformprojektes, einer Innovation usw.) voranzubringen. Kuhlmann sieht nun Ansätze, diese Gegenüberstellung aufzulösen und die divergierenden Interessen beteiligter Akteure in einen Verhandlungsprozess einzubringen und so aufzubrechen. Der Evaluator tritt nun als Moderator und als „facilitator“ von Lernprozessen auf. Evaluation ist dann nicht mehr Leistungsmessung, sondern fungiert als Lernmedium.

Anthony F.J. van Raan plädiert in seinem Beitrag nicht für einen Ersatz der qualitativen und subjektiven Urteile im „peer review“, sondern für eine Ergänzung mit bibliometrischen Verfahren, die sich vor allem für höher aggregierte Zitationsdaten eignen, also für Forschergruppen oder ganze Institute: **The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments.** Er demonstriert einen neuen Auswertungsansatz, kommentiert den Einsatz auf sozialwissenschaftlichen Feldern und schlägt im letzten Teil einen neuen Darstellungsansatz („structure mapping“) vor, um langfristige Entwicklungen in einem Forschungsfeld aufzuzeigen. Bei wissenschaftlichen Publikationen ist nicht nur entscheidend, wer wie viel in welchen Organen publiziert, sondern

auch wer wie häufig wo zitiert wird und damit Eingang in die wissenschaftliche Kommunikation gefunden hat. Bei diesen Zitierungen, zu denen es Datenbanken gibt, wird aber, das ist dem Laien vermutlich nicht klar, nicht erfasst, *wie* zitiert wird (also z. B. bestätigend, ablehnend, eine Hypothese differenzierend usw.), sondern nur „dass“. Natürlich kann man, z. B. in Meta-Analysen zu bestimmten Forschungsfeldern, die Art der Verwertung genau erfassen, aber das geht nicht ohne eigene kenntnisreiche Lektüre. In den Datenbanken, mit denen nachfolgend gearbeitet wird, ist diese Verwertungsart nicht erfasst.

Am Beispiel der Publikationszahlen eines „German medical research institute“ der Molekularbiologie im Zeitraum von 1992 bis 2000 demonstriert van Raan einen weiterentwickelten Ansatz, dessen Logik darin besteht, mit einer sukzessiv erweiterten Datenbasis zu arbeiten, so dass die Vergleiche auf einem immer breiteren Grund stehen. Die Anzahl an Zitierungen, die ein Artikel in einem Fünfjahreszeitraum z. B. erhält (citation per paper, CPP), kann z. B. ins Verhältnis gesetzt werden zu einer durchschnittlichen Zitierungsrate bei *allen Artikeln* jener Journale, in denen das Institut publizierte, oder zu einer Rate in allen Zeitschriften eines *bestimmten Gebietes* (weltweit). Diese Beispielrechnungen werden auf der Basis der Zitierungsdatenbanken des ISI, des Institute for Scientific Information, durchgeführt; man weiß aber, dass diese einen amerikanischen Bias haben.

Van Raan betont selbst, dass sein methodischer Ansatz vor allem bei den Natur- und den Lebenswissenschaften anwendbar ist, weniger in den angewandten Wissenschaften, noch weniger in den Sozial- und Geisteswissenschaften. Hier herrschten doch andere und vor allem in den einzelnen Gebieten auch unterschiedliche Publikationskulturen. Doch möchte er nicht gelten lassen, dass Sozialwissenschaften für bibliometrische Analysen gar nicht zugänglich seien. Auch hier plädiert er für forschungsfeldbezogene Analysen mit möglichst hoch aggregierten Datensätzen – und in längeren Zeiträumen (dann eher 5 bis 6 Jahre als 2 bis 3 wie in den Naturwissenschaften).

Mit einem Datensatz von mehreren Tausend Artikeln aus dem Gebiet der Mikroelektronik demonstriert van Raan eine neuartige Dar-

stellungsmethode, die die Größe und thematische Nähe von einzelnen Forschungsfeldern zeigt (z. B. ein Cluster um Electronic structures, Materials und Liquids/Solids Structures). Grundlage der Bildung der Ähnlichkeitsmaße sind Titel und Abstract der einzelnen Artikel und deren Übereinstimmungen bei den verwendeten Begriffen. Ob und inwieweit damit aber Wissens- und Wissenschaftsstrukturen („processes of knowledge dissemination“, so im Abstract) erfasst werden, muss offen bleiben. Selbstverständlich sind Titel und Abstract eines Zeitschriftenartikels wichtige Elemente im wissenschaftlichen Kommunikationssystem, aber dieses umfasst eben weit mehr.

Im dritten Beitrag dieses Blocks geht es um die **Evaluation des Zentrums für Technologiefolgen-Abschätzung TA-SWISS 2002**, und die Darstellung ist das Ergebnis einer gemeinsamen Schreibanstrengung der Evaluierten (Walter Grossenbacher-Mansuy, Sergio Bellucci für TA-SWISS) und der Evaluierenden (Sami Kanaan für Evaluanda). Dies hat seinen Grund, denn im methodischen Ansatz wurde eine Selbstevaluation mit einer Fremdevaluation einer externen Instanz (eben Evaluanda) kombiniert. Dabei umfasste diese Selbstevaluation nicht nur das Sammeln und Bereitstellen der nötigen Unterlagen, sondern wirklich die Formulierung einer kritischen Selbsteinschätzung hinsichtlich der Stärken, Schwächen, Chancen und Risiken. Die ganze Evaluation wurde im Übrigen vom Leitungsausschuss der Einrichtung veranlasst. Um dies besser einordnen zu können, bat die Redaktion TA-SWISS um ein Organigramm sowie einige Erläuterungen (siehe die Abbildung).

Für die Ausarbeitung des Evaluationsdesigns wurden drei Ebenen unterschieden, die man in drei Vergleichsschritte gruppieren kann: a) gesellschaftlich/politischer Bedarf und der institutionelle Auftrag; b) dieser Auftrag und die internen Bearbeitungsrountinen; und c) die internen Prozesse und deren Ergebnisse, die an bestimmte Adressaten gehen (auch Parlament), aber auch in die allgemeine Öffentlichkeit hineinwirken sollen. Jedes Teilprogramm wäre genug Stoff für eine Evaluation gewesen, aber man wollte partout das ganze Programm; entsprechend gestaltete sich der Aufwand. Ob sich auch die Wirkungen auf der politischen und gesellschaftlichen Ebene eingestellt haben,

ist - da ist der Bericht sehr nüchtern – offen und bestenfalls qualitativ festzustellen.

Diese Evaluation spielte sich dann in einem Kontext knapper Kassen ab, der den positiven Ertrag (Mittelaufstockung wurde empfohlen) konterkarierte und praktisch zunichte machte. Es gibt also (nicht nur in diesem Fall) eine Asymmetrie der Evaluationswirkung. Eine Mittelaufstockung aufgrund der positiven Evaluation ist unwahrscheinlich, dagegen hätte ein Auslassen der Evaluation mit hoher Wahrscheinlichkeit Mittelverluste bedeutet.

Die beiden zentral stehenden Beiträge sind der Evaluationspraxis des Wissenschaftsrates gewidmet. Der Wissenschaftsrat ist, wenn es um die Evaluation der Hochschulforschung, aber auch der außeruniversitären Forschung geht, sicher die erste Adresse. Es wurden mittlerweile alle „Säulen“ des deutschen Forschungssystems evaluiert und in ihrer Grundausrichtung und wechselseitigen Profilierung für gut befunden (dies gilt auch für die HGF; vgl. WR 2001b). Insbesondere bei den Instituten der Blauen Liste war der Evaluationsansatz des WR herausgefordert, weil es dort Institute mit Beratungsaufgaben oder sogar Serviceaufgaben gibt. Die Redaktion bzw. die Institutsleitung hatte darum gebeten, dass **Reinhard Hüttl** in seinem Beitrag diesen Gesichtspunkt aufgreift: **Evaluation politikberatender Forschungsinstitute durch den Wissenschaftsrat – Kriterien und Erfahrungen.**

Die einzelnen Etappen und Überlegungen werden nachgezeichnet. Aus Sicht einer einführenden Charakterisierung sind es vor allem zwei Aussagen, die u. E. festzuhalten sind, (1) der unter den heterogenen Aufgabenstellungen (vor allem der Institute der Blauen Liste) entstehende Bedarf nach einer Weiterentwicklung und Revision des Evaluationsansatzes des Wissenschaftsrates, und (2) der immer wieder bestärkte Grundsatz, dass die Basis der Evaluation die „Qualität der wissenschaftlichen Arbeit“ darstellen muss. Dies umschließt auch „wissenschaftliche Politikberatung“, auch sie muss „fachlich qualifiziert und auf der Grundlage guter wissenschaftlicher Arbeit“ erfolgen (Zitat aus der abschließenden Stellungnahme zu den Evaluationen der Blauen Liste).

Vermutlich liest man diesen Hinweis so, dass die Beratung auf der Grundlage eigener Forschungsleistungen zum Gegenstand der Be-

ratung zu erfolgen habe – wer zu Arbeitsmarktpolitik berät, sollte über arbeitsmarkt- und arbeitsmarktpolitische Fragestellungen forschen; wer zu Klimapolitik berät, sollte selbst Klimaforschung betreiben. Liest man in der schon mehrfach erwähnten Systemevaluation der Blauen Liste nach, dann wird eine solche objektwissenschaftliche Interpretation nahe gelegt; liest man in den „Verfahrensgrundsätzen“ nach (WR 2002), die auch Hüttl zitiert, dann wird bei der „Bewertung von Beratungsleistungen“ deutlich, dass auch eine prozesswissenschaftliche Ausdeutung möglich ist, denn natürlich kann man auch den Beratungsprozess professionalisieren, methodisieren und in diesem Sinne verwissenschaftlichen.²

Obwohl vielleicht von Hüttl und dem Evaluationsausschuss so nicht gemeint, könnte man aber im Sinne einer Wiedereröffnung der Diskussion beide Pole offen halten. Eine solche Wiedereröffnung und auch Weiterführung der Diskussion versuchen im dann folgenden Beitrag **Hans-Jochen Luhmann** und **Thomas Langrock** vom Wuppertal Institut. Man weiß, dass der Wissenschaftsrat das WI im allgemeinen und die Abteilung Klimapolitik im besonderen nicht positiv evaluiert hat, und der Streit über das Votum ist ja in den öffentlichen Medien geführt worden. Hier ist nicht das Forum, diese Kontroverse neu aufzulegen. Der Redaktion war klar, dass es schwierig werden würde (und es vielleicht auch gar nicht opportun sei), jemanden aus dem Kreis der Betroffenen zur Feder greifen zu lassen, aber die Diskussion muss ja weitergeführt werden. Nach Meinung der Redaktion haben Luhmann und Langrock das Menschenmögliche getan, sich der rhetorischen Reflexe von Evaluations-Betroffenen zu entledigen und sich auf die analytische Substanz zu konzentrieren, um die der Streit geht, nämlich das angemessene Wissenschaftsverständnis: **Was ist wissenschaftliche Politikberatung?**

Luhmann und Langrock greifen dabei auf die Aufgabenstellung des Instituts zurück, in dem sie ein spezifisches Wissenschaftsverständnis und einen Diskursauftrag erkennen, rekonstruieren das Wissenschaftsverständnis wie in der WI-Evaluation durch den Wissenschaftsrat ausgedrückt und kommen naheliegenderweise zu anderen Konsequenzen als der Wissenschaftsrat: Dieser empfahl, eben die wissenschaftlichen Grundlagen der Beratung

zu pflegen (in Übereinstimmung mit den o. g. Grundsätzen), während die WI-Vertreter dafür plädieren, den – wie wir es ausdrücken würden – Beratungs- und Diskursauftrag zu professionalisieren. Wir finden: Darüber kann (und sollte) man streiten. Bei der Aufarbeitung der Evaluationserfahrungen der Blauen Liste wären im übrigen auch die Befragungsergebnisse auszuwerten, die Röbbcke und Simon (2001) zusammengetragen haben.

Die drei Beiträge aus dem Kreis der HGF wurden eingangs schon als Erfahrungsberichte charakterisiert. Im Rahmen der lebhaften Korrespondenz während der Vorbereitung des Heftes fiel einmal die Bemerkung, dass es leichter sei, einen wissenschaftlichen Artikel als einen solchen Erfahrungsbericht zu verfassen. Das sieht die Redaktion ähnlich. Deshalb sei den Autoren und der Autorin besonders gedankt: **Manfred Popp** (Vorstandsvorsitzender des Forschungszentrums Karlsruhe), der **Erste Schritte in die Programmorientierte Förderung** kommentiert; **Henning Möller**, im Forschungszentrum Karlsruhe Leiter der Stabsabteilung Planung, Außenbeziehungen und Erfolgskontrolle, der **Zur Entwicklungsgeschichte von Steuerung und Erfolgskontrolle in der Helmholtz-Gemeinschaft** Stellung nimmt, sowie **Susanne Schultz-Hector** vom GSF-Forschungszentrum, die die Erfahrungen aus der Evaluation der Gesundheitsforschung in der HGF berichtet: **Begutachtung des Helmholtz-Forschungsbereiches Gesundheit: Ein Erfahrungsbericht**.

Man versteht aber diese Beiträge nicht – und die Leser und Leserinnen seien gewarnt – wenn man sich nicht diese neuen (und nicht einfachen) HGF-Strukturen vor Augen führt und eine Ahnung von den verzweigten Entscheidungsprozeduren entwickelt. Deshalb war es unabdingbar, wenigstens einen Teil dieser Strukturen in einer eingelegten Doppelseite darzustellen und zu erläutern.

Als zentral stellen wir folgende Merkmale heraus: (a) Die Finanzierung der Zentren erfolgt nicht mehr institutionell, sondern gemäß zentrenübergreifenden Programmen; um die Ressourcenanteile bewerben sich die Einrichtungen in der Regel in gemeinsamen Programmen; (b) neben die bisher schon regelmäßig stattfindenden zentren-bezogenen Evaluationen treten weitere in einem 5-Jahreszyklus; diese

jetzt begonnenen Evaluationen sind „ex ante“-Evaluationen, d. h. die vorgelegten Programme sind F+E-Entwürfe für die kommenden Jahre; (c) die Begutachtungen ziehen im HGF-Senat und im Ausschuss der Zuwendungsgeber konkrete Finanzierungsentscheidungen nach sich.

Aus den drei Beiträgen sei im Folgenden nur je ein Aspekt hervorgehoben: Im Beitrag von **Popp** die Sorge, dass sich die Gutachter zu eng auf die einzelnen Programme fokussieren und wichtige Verbindungen zu anderen Stärken der Zentren gar nicht mehr wahrnehmen. Aus dem Beitrag von **Möller**, der bis auf die Anfänge der Erfolgskontrolle der Zentren zurückblickt, die Warnung, dass auch das ausgefeilteste Controlling-Instrumentarium wenig ausrichten kann, wenn nicht ein gewisser Konsens hinsichtlich der einzuschlagenden Strategien vorhanden ist. Im Beitrag von **Schultz-Hector** werden die Entscheidungsetappen nochmals skizziert sowie Grundlagen und Verfahren der Beratung der HGF-Gesundforschung dargelegt; sie betont in ihrem persönlichen Statement u. a. die neuen Kooperationen und gedanklichen Anknüpfungspunkte, die sich in den Programmberatungen ergeben haben.

In diesen neuartigen Evaluationen hatte der HGF-Senat die Gutachter bestellt, der Wissenschaftsrat war hier nicht gefordert. Aber es soll, evaluationsmethodisch konsequent, nach Abschluss der Begutachtungen das Begutachtungsverfahren selbst einer kritischen Würdigung durch den WR unterzogen werden.

Indikatoren für eine diskursive Evaluation transdisziplinärer Forschung – dieser Titel des Beitrages von **Matthias Bergmann** umschreibt das Ziel, wo das Projekt, in dem diese Indikatoren erarbeitet und entwickelt werden sollen, ankommen will. Dieses Projekt steht noch am Anfang und insofern wird dankbar verzeichnet, dass schon jetzt ein Einblick in die Projektwerkstatt gegeben wird. Das Projekt ist eingebettet in den Forschungsverbund sozial-ökologischer Forschung, nennt sich „Evalunet“ und will ein „Evaluationsnetzwerk für transdisziplinäre Forschung“ etablieren. Konkret soll die Forschungspraxis von sechs transdisziplinär ansetzenden Forschungsprojekten auf Erfolgsfaktoren und eben die genannten Indikatoren hin ausgewertet werden, so dass am Ende so etwas wie ein Regelwerk für angemessene Evaluationsverfahren steht. Transdisziplinäre

Projekte, das ist die These, müssen anders evaluiert werden. Zwei dieser Projekte wurden schon ausgewertet, und der Beitrag stellt erste Einsichten dar.

Wie wird vorgegangen und warum „diskursive“ Evaluation? Die Projektberichte werden ausgewertet, zusätzliches Material (z. B. Sitzungsprotokolle über kritische Fragen in der Konstitutionsphase eines Projektes) analysiert, vom Evalunet-Projekt werden zusätzliche Befragungen durchgeführt und – hier kommt das diskursive Element hinein – in einer gemeinsamen Runde, mit extern berufenen und für das jeweilige Projekt spezifisch zusammengestellten Experten, werden in einem dreitägigen Workshop die ganzen Fragen, Erfahrungen, Irrungen und Lösungsansätze diskutiert, und damit auch, so jedenfalls haben wir das verstanden, einer theoretisierenden Rekonstruktion unterzogen. Und diese Rekonstruktion ist nötig, denn bloße Deskription würde ja für das Regelwerk nicht hinreichen. Aber die Stoßrichtung ist klar eine formative, wie einleitend schon charakterisiert: Evaluation als Lernmedium. Wir sind gespannt auf die weiteren Resultate.

Die Redaktion dankt der Autorin und allen Autoren, die zu diesem Themenschwerpunkt beigetragen haben, und dies in einer Zeitschrift, die eher in einem Feld vorgelagerter Diskussionen operiert und insofern nicht als eine „normale“ Zeitschrift anzusehen ist und im Science Citation Index nicht geführt wird. Aber über Zitierungsumwege kann man trotzdem dort ankommen; eine erste Recherche, „quick and dirty“, ergab immerhin zwei Treffer. Aber der eine Kollege ist schon lange nicht mehr da und der zweite auch nicht, jedenfalls nicht mit diesem Namen. Man sieht: Vorsicht und Sorgfalt sind angebracht!

(Bernd Wingert und Reinhard Coenen, ITAS)

Anmerkung

- 1) Der Wissenschaftsrat ist eine gemeinsam von Bund und Ländern eingesetzte Einrichtung zur wissenschaftlichen Beratung von Bund und Ländern; vgl. <http://www.wissenschaftsrat.de>.
- 2) Es ist klar, dass ab einer bestimmten Kontingenz von gesellschaftlicher Problemstellung, organisiertem Forschungspotential und darauf basierenden Entscheidungen eine schlichte Trennung in Objektwissen und Prozesswissen nicht mehr hinreicht. Die Frage ist auch, ob mit neuartigen Problemstellungen ein neuartiger Forschungstyp entsteht („mode 2“), der mit der überkommenen Begrifflichkeit nicht mehr gefasst werden kann. Diese Diskussion wurde schon 1999 in Heft 3/4 ein Stückweit geführt.

Literatur

Hellstern, G.; Wollmann, H., 1980: Evaluierung in der öffentlichen Verwaltung – Zweck und Anwendungsfelder. In: Verwaltung und Fortbildung, S. 61ff

Lawrence, P.A., 2003: The politics of publication. Nature 422, S. 259-261

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

Stockmann, R., 2000: Evaluation in Deutschland. In: Stockmann, R. (Hrsg.): Evaluationsforschung. Grundlagen und ausgewählte Forschungsfelder. Opladen: Leske + Budrich, S. 11-49

Wissenschaftsrat, 2001a: Systemevaluation der Blauen Liste – Stellungnahme des Wissenschaftsrates zum Abschluss der Bewertung der Einrichtungen der Blauen Liste. Drs. 4703/00; 19.1.2001 (siehe auch: <http://www.wissenschaftsrat.de>)

Wissenschaftsrat, 2001b: Systemevaluation der HGF – Stellungnahme des Wissenschaftsrates zur Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren. Drs. 4755/01; 19.1.2001 (siehe auch: <http://www.wissenschaftsrat.de>)

Wissenschaftsrat, 2002: Aufgaben, Kriterien und Verfahren des Evaluationsausschusses des Wissenschaftsrates. Drs. 5375/02; 12.7.2002 (siehe auch: <http://www.wissenschaftsrat.de>)

«