Leistungsmessung oder Lernmedium? Evaluation in der Forschungs- und Innovationspolitik¹

von Stefan Kuhlmann, Fraunhofer Institut für Systemtechnik und Innovationsforschung (ISI)

Politik und Gesellschaft verlangen zusehends Rechenschaft über Leistungen. Qualität und Nutzen von staatlich geförderter Wissenschaft und Forschung. Forschungsprojekte und Förderungsprogramme sowie Forschungsinstitutionen werden evaluiert. Neue Autonomiespielräume der Hochschulen gehen einher mit einer Zuweisung von Forschungsmitteln auf der Grundlage von Evaluation. Forschungsmanagement und Politik haben einen wachsenden Bedarf an evaluativer Information. Hier stellt sich die Frage, ob solche Information vor allem zur "Verwaltung des Mangels" an Forschungsmitteln oder auch zur intelligenten, kreativen Weiterentwicklung des Forschungsund Innovationssystems und seiner Elemente verwendet wird. Der Beitrag enthält eine knappe Darstellung der Entwicklungslinien der staatlichen Forschungs- und Innovationsförderung, bietet danach eine Übersicht der wichtigsten Konzepte, Anwendungsbereiche und Methoden evaluativer Verfahren, diskutiert sodann die Verwendung von Evaluation im Prozess politischer Entscheidungsfindung und schließt mit einem Ausblick auf methodisch angereicherte akteurs- und prozessorientierte d. h. lernorientierte – Evaluationsverfahren.

1 Einführung

Die Institutionen der wissenschaftlichen Forschung und ihre Kommunikationen untereinander sind als selbstreferentielles Systeme beschrieben worden, dessen Leistungsorientierungen und Qualitätsmaßstäbe sich deutlich von denen anderer gesellschaftlicher Subsysteme unterscheiden (Luhmann 1990). Ein Teil der heute verwendeten Evaluationsverfahren – vor allem die Praktiken der Bewertung wissenschaftlicher Leistungen, Projekte und Publikationen durch Fachkollegen (peer review) – wurzelt in dieser Selbstbezüglichkeit. Doch sie hat

ihre Grenzen: Politik und Gesellschaft verlangen seit den 1990er Jahren verstärkt – Rechenschaft über Leistungen, Qualität und Nutzen von staatlich geförderter Wissenschaft und Forschung, teils weil staatliche Mittel knapp sind und möglichst "effektiv" verwendet werden sollen, teils weil Politik und Öffentlichkeit eine gewisse Skepsis gegenüber diesem selbstbezüglichen Wirken hegen. Große Forschungsprojekübergreifende Forschungsförderungsprogramme und Forschungsinstitutionen wurden zusehends zum Gegenstand von Evaluationsverfahren. Wachsende Autonomiespielräume der Hochschulen durch Einführung von Globalhaushalten gehen einher mit leistungsabhängiger Ressourcenallokation. Das in Gang gesetzte Prinzip des Wettbewerbs zwischen Forschungseinrichtungen erfordert eine verstärkte Profilbildung durch gezieltes Wissenschaftsmanagement. Die Europäische Kommission arbeitet gegenwärtig an entsprechenden Indikatoren für den künftigen "europäischen Forschungsraum" (Noyons, Schmoch 2003). In der Summe bewirkt dies in Forschung, Management und Politik einen wachsenden Bedarf an evaluativer Information. Hier stellt sich die Frage, ob solche Information vor allem zur "Verwaltung des Mangels" an Forschungsmitteln oder auch zur intelligenten, kreativen Weiterentwicklung des Forschungssystems und seiner Elemente verwendet wird.

2 Entwicklungslinien der Forschungs- und Innovationsförderung

Wenn hier von Forschungs- und Innovationspolitik die Rede ist, dann sind damit alle politischen Initiativen zur Gestaltung des Forschungsund Innovationssystems gemeint (Larédo, Mustar 2001), dies schließt seit den 1970er Jahren Versuche ein, das Innovationsverhalten industrieller Unternehmen positiv zu beeinflussen. Das deutsche System gilt im internationalen Vergleich als relativ gut entwickelt: 2000 waren knapp 480.000 Personen (Vollzeitäquivalente) mit Forschung und Entwicklung (FuE) beschäftigt; die Gesamtausgaben für FuE beliefen sich auf knapp 50,1 Mrd. Euro, das entspricht 2,48 % des Brutto-Inlandprodukts (Zahlen nach BMBF 2002). Die Forschungsinfrastruktur ist vergleichsweise differenziert:

- Die *Industrie* realisiert den größten Anteil von FuE in Deutschland (306.700 Mitarbeiter): 1999 investierte sie knapp 33,6 Mrd. Euro überwiegend in angewandte Forschung und experimentelle Entwicklung.
- Auf die Hochschuleinrichtungen (345 staatliche bzw. staatlich anerkannte Hochschulen, darunter 116 Universitäten und vergleichbare Einrichtungen; 101.500 FuE-Mitarbeiter) fällt der zweitgrößte Anteil der Forschungsausgaben (8,1 Mrd. Euro). Sie konzentrieren sich auf Grundlagen- und langfristig anwendungsorientierte Forschung, finanziert durch die Bundesländer sowie durch die Deutsche Forschungsgemeinschaft (DFG; etwa 1,2 Mrd. Euro). Im Verlaufe der 1980er Jahre ist der Anteil industrieller Forschungsaufträge an den Forschungsbudgets einzelner Universitäten, insbesondere technische, signifikant gewachsen (auf durchschnittlich 11,3 %).
- Die 15 "Großforschungseinrichtungen" der Helmholtz-Gemeinschaft (HGF) des Bundes (21.500 FuE-Mitarbeiter; 2,1 Mrd. Euro) leisten vor allem langfristig orientierte Forschung, die als risikoreich gilt, hohe Kosten verursacht (Anlagen) und große Forschungsteams benötigt. Daneben betreiben verschiedene Ministerien des Bundes sog. Ressort-Forschungseinrichtungen zur wissenschaftlich-technischen Unterstützung ihrer Aufgaben.
- Die Institute der Max-Planck-Gesellschaft (MPG; 9.200 FuE-Mitarbeiter; 1 Mrd. Euro) konzentrieren sich auf ausgewählte Felder der Grundlagenforschung in den Naturwissenschaften und Geisteswissenschaften.
- Die Institute der Fraunhofer-Gesellschaft (FhG; 2001: 12.000 FuE-Mitarbeiter; 992 Mio. Euro, davon 853 Mio. Vertragsforschung) sollen die praktische Verwendung wissenschaftlicher Erkenntnisse durch langfristig anwendungsorientierte und angewandte Forschung fördern. Die Fraunhofer-Gesellschaft führt in erster Linie Vorhaben der Auftragsforschung durch, die teils von der Industrie und teils von staatlichen Stellen finanziert wird. Keine andere Forschungsorganisation in Deutschland ist in den vergangenen 20 Jahren so schnell gewachsen.
- Die Institute der *Wissenschaftsgemeinschaft G.W. Leibniz* (WGL; früher "Blaue Liste";

- 10.000 FuE-Mitarbeiter; 811 Mio. Euro) schließlich bilden eine Restkategorie, deren Gemeinsamkeit vor allem darin besteht, dass diese Einrichtungen von Bund und Ländern institutionell gefördert werden.
- Die Forschungseinrichtungen der Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF) führen vor allem angewandte Forschung und experimentelle Entwicklung für sektor-spezifische Bedürfnisse industrieller Unternehmen durch.

2000 gab der Bund insgesamt 8,4 Mrd. Euro für Forschung und Entwicklung aus. Das Spektrum von Instrumenten der staatlichen Forschungsund Innovationspolitik ist heute weit ausdifferenziert und reicht von der *institutionellen* Förderung von Forschungseinrichtungen (45,6 %)
über verschiedene Formen finanzieller Anreize
(*Programme*; 40,5 %) zur Durchführung von
Forschung und experimenteller Entwicklung in
öffentlichen oder industriellen Forschungslaboratorien bis zur Gestaltung einer "innovationsorientierten" Infrastruktur.

Moderne Forschungs- und Innovationssysteme standen immer und stehen gegenwärtig ganz besonders unter dem *Druck der Anpassung* an veränderte Umwelten:

- Die Entwicklung von Hochtechnologieprodukten hat in den vergangenen Jahren eine signifikante *Beschleunigung* erfahren; die Hervorbringung einer neuen Idee und ihre Kommerzialisierung am Markt folgen heute so schnell aufeinander wie nie.
- In vielen Gebieten wächst die "Wissenschaftsbindung" des Innovationsgeschehens; Grundlagenforschung (etwa molekularbiologische) wird auf die Entwicklung marktfähiger Produkte gerichtet durchgeführt (Meyer-Krahmer, Schmoch 1998).
- Dabei erfordern komplexe Produkte und Prozesse eine dichtere Verflechtung und "Fusion" (Kodama 1995) heterogener Technikentwicklungen, wodurch traditionelle Grenzziehungen zwischen Wissensund Technikgebieten verschwinden und interdisziplinäre Kompetenzen im Innovationsprozess an Bedeutung gewinnen, neue disziplinäre Cluster entstehen und alte verschwinden.
- Die genannten Trends steigern insgesamt die Erfordernis der *Kooperation* von Inno

vationsakteuren: Mit wachsender Komplexität des erforderlichen Wissens sind isolierte Akteure immer weniger in der Lage, dieses ohne externe Unterstützung hinreichend zu beherrschen². Kooperation und internationale Netzwerke (Callon 1992) gehören mittlerweile zum alltäglichen Innovationsgeschehen.

- Nur scheinbar im Widerspruch hierzu steht die Herausbildung von besonders starken, an einen *bestimmten Standort* gebundenen, thematisch fokussierten Forschungs- und Innovationskapazitäten, die weltweit im Wettbewerb mit nur ein oder zwei anderen Standorten stehen, wobei sich die Rangfolge durchaus ändern kann (z. B. Automobilbau in Deutschland; gentechnologische Pharmazeutik in den USA; Bürotechnik in Japan) (Meyer-Krahmer 1999).
- Einer neuen Dynamik ist schließlich auch das Verhältnis von kodifiziertem und nichtkodifiziertem Wissen im Innovationsprozess ausgesetzt. Computernetzwerke beschleunigen den Umlauf kodifizierten Wissens, das nur dann effektiv nutzbar ist, wenn es re-kontextualisiert werden kann, wenn also die *Lernfähigkeit der Innovationsakteure* wächst (Lundvall, Borrás 1998).

Staatliche Politik reagiert auf diese Trends in Wissenschaft und Innovation, indem sie mit ihrer Förderpolitik strukturellen Wandel in der "Forschungs- und Innovationslandschaft" zu bewirken sucht und Modernisierungsdruck auf Forschungsinstitutionen ausübt: So werden seit Mitte der 1990er Jahre viele Förderprogramme als thematische Wettbewerbe ausgeschrieben, die auch einen strukturellen Wandel bewirken sollen. Partnerschaften verschiedenster Einrichtungen (Forschungseinrichtungen, Universitäten, private Firmen, Technologiezentren, Weiterbildungseinrichtungen etc.) zielen mit einem ganzen Bündel aufeinander abgestimmter Maßnahmen auf eine Effektivierung ganzer Innovationssysteme.

Die veränderten Anforderungen an Institutionen wurden im Rahmen der von einer internationalen Expertengruppe durchgeführten "Systemevaluation" der MPG und der DFG prägnant zusammengefasst. Die Gutachter forderten unter anderem eine Lockerung der vorherrschenden disziplinären Orientierung, die Entwicklung beweglicher und leistungsfähiger

Organisationsformen für eine temporäre Zusammenarbeit verschiedener Disziplinen und Gruppen in problemorientierten Forschungsfeldern, wirksame Verfahren zur Qualitätssicherung unter externer Beteiligung, eine verbesserte Zusammenarbeit von Hochschulen und außeruniversitären Forschungseinrichtungen, die Förderung von Institutionen übergreifenden Forschungszentren sowie eine verstärkte internationale Vernetzung der Einrichtungen (Internationale Kommission 1999).

3 Evaluationsverfahren

3.1 Konzepte und Anwendungsgebiete

In historischer Perspektive kann man in Deutschland wie auch in anderen industrialisierten Ländern³ wissenschafts*interne* und *externe* Ansätze als zwei heterogene Entwicklungslinien der Evaluation von Wissenschaft sowie Forschungs- und Innovationsförderung unterscheiden. Sie lassen sich als ein Drei-Schalenmodell darstellen (vgl. Kuhlmann 2003a):

Erste Schale – Individuelle Forschungsleistungen: Den "Kern" bilden peer review-Verfahren und später zusätzlich Verfahren zur Messung der Forschungsleistung einzelner Forscher und Gruppen (Bibliometrie etc.) als wissenschaftsinterne Instrumente für die Entscheidung über die Allokation von Fördermitteln in Forschungseinrichtungen (Guston 2003; Hornbostel 1997; Daniel 1993; van Raan 1988). Das Peer Review-Verfahren ist das vorherrschende Evaluationsinstrument der Deutschen Forschungsgemeinschaft (DFG). Förderanträge werden von Peers bewertet, die im Vier-Jahres-Rhythmus von der gesamten wissenschaftlichen Gemeinschaft gewählt werden (Neidhardt 1988). Jeder Gutachter ist angehalten, den Antrag allein auf der Grundlage seiner wissenschaftlichen Qualität zu beurteilen. Kürzlich haben Senat und Hauptausschuss der DFG beschlossen, dieses bewährte System ("bottom-up") um strategieorientierte, interdisziplinäre und internationale Komponenten zu ergänzen (DFG 2002).

Zweite Schale – Programme: Um diesen Kern der Peer Review-Verfahren legt sich eine Schale von Evaluationsstudien im Sinne von impact analyses forschungs- und innovati onspolitischer Programme, die im Unterschied zur institutionellen Förderung gezielte Gestaltungsabsichten verfolgen. In Deutschland zeigte sich insbesondere das Bundesforschungsministerium aktiv; zwischen 1985 und 1993 ließ man etwa 50 größere Evaluationsstudien anfertigen (Kuhlmann/Holland 1995a), in den folgenden Jahren nicht weniger. Solche Studien schließen an die amerikanische Wirkungsforschung und policy analysis an ("policy-analytic movement"; Schön/Rein 1994; Übersichten bei Widmer 1996; Patton 1997; Kuhlmann 1998). Dieses Evaluationskonzept kann als wissenschaftsextern bezeichnet werden, denn es wird überwiegend vom politischadministrativen System in Gang gesetzt, um die Erreichung politisch gesetzter Ziele zu prüfen. Als Evaluatoren treten dabei üblicherweise unabhängige Forschungsinstitute im Auftrag der Administration auf; sie wenden ein relativ breites Spektrum von Evaluationskonzepten und -instrumenten an (Shapira, Kuhlmann 2003) und sind seit den 1990er Jahren auch professionell in einer "Deutschen Gesellschaft für Evaluation (DeGEval)" organisiert.

Dritte Schale - Institutionen: Hier geht es um die Leistungsfähigkeit von Forschungsinstitutionen oder größeren Wissenschaftsgebieten, wie sie u. a. der Wissenschaftsrat (WR) durchführt; nach der deutschen Vereinigung nahmen die WR-Evaluationen gestaltenden Einfluss auf die neue Struktur der Forschungslandschaft Ostdeutschlands (Block, Krull 1990). Seit den 1990er Jahren wurden institutionelle Evaluationen immer häufiger durchgeführt (Kuhlmann, Holland 1995b). Im Frühjahr 1999 schloss eine internationale Kommission die "Systemevaluation" der DFG und der MPG ab (Internationale Kommission 1999). Zeitlich parallel wurde auch eine Systemevaluation der FhG durchgeführt (vgl. Evaluierungskommission 1998). Eine aufwändige Evaluation der Einrichtungen der WGL wurde 2000 abgeschlossen (Wissenschaftsrat 2000; Röbbecke, Simon 2001) sowie 2001 Systemevaluationen der Großforschungseinrichtungen der HGF (Wissenschaftsrat 2001) und der Institutionen der "industriellen Gemeinschaftsforschung" (Kommission Systemevaluation 2001). Bei der Evaluation von Forschungsinstitutionen werden zunehmend schaftsinterne und -externe Fragestellungen und Informationsquellen kombiniert. Parallel dazu widmete auch die Privatwirtschaft der Effizienz, Effektivität und strategischen Ausrichtung ihrer Laboratorien erhöhte Aufmerksamkeit (z. B. Brockhoff 1999; Bürgel et al. 1996).

Zusammenfassend: Die deutsche Evaluationspraxis ist einerseits *entwickelt* hinsichtlich des hohen Maßes an Selbstorganisation bei wissenschaftsinternen Verfahren, die Konsens und Verpflichtung zwischen den Forschern fördern, sowie *dynamisch* wegen vielfältiger neuer Evaluationsansätze und Strategieentwicklungen in allen Institutionen, sie ist jedoch auch *unsystematisch* und *zersplittert*, weil die verschiedenen institutionell orientierten Evaluationsbemühungen bisher nur wenig aufeinander abgestimmt werden konnten.

Weiterhin ist festzustellen, dass sich viele Evaluationsinitiativen (interne und externe Ansätze) nicht mehr allein auf die Feststellung der Qualität einzelner Forschungsleistungen beschränken. Zusehends soll evaluative Information auch zur Erhellung strukturbildender sowie sozioökonomischer Effekte von Forschung und Forschungsförderung beitragen (vor allem aus der Perspektive der Europäischen Union: Airaghi et al. 1999; Polt et al. 2002; Georghiou et al. 2003). Ein gemeinsames Raster von Indikatoren zur Erfassung und Analyse solcher Effekte wurde bisher jedoch noch nicht verabredet.

3.2 Methoden

Wir kennen heute vielfältige Methoden zur Feststellung erzielter oder erzielbarer Wirkungen von Forschung und Forschungsförderung. Die wichtigsten sind der Vorher-/Nachher-Vergleich, der Kontroll- oder Vergleichsgruppenansatz, sowie qualitative Analysen (u. a. Plausibilitätsüberprüfungen, Schätzurteile). Sie können mit unterschiedlichen Indikatoren (finanzieller Aufwand für Forschung und Entwicklung, Patente, ökonomische, soziale, technische Kenngrößen, Veröffentlichungen, Zitate etc.), Datensammlungsverfahren (Statistiken, Fragebögen, Interviews, Fallstudien, Panel etc.) und Datenanalyseverfahren (ökonometrische Modelle, Cost-/Benefit-Analysen, andere statistische Verfahren, Technometrie, Bibliometrie, Peer Review) einzeln oder kombiniert verwendet werden (Meyer-Krahmer 1989; Bozeman, Melkers 1993; Polt et al. 2002). Bei allen notwendigen Bemühungen um objektivierende Verfahren und geeignete Indikatoren muss jedoch davor gewarnt werden, quantitative Indikatoren allein als hinreichend für die Evaluation von Forschungs- und Innovationsförderung zu betrachten. Der verständliche Wunsch nach einem standardisiert anwendbaren "Indikatoren-Werkzeugkasten" ist bei der Tendenz zur Verfolgung komplexer politischer Zielsetzungen nicht erfüllbar (Airaghi et al. 1999).

Ein aus der Perspektive der Evaluationsforschung akzeptables, methodisch abgerundetes Konzept einer ex post-Evaluation politischer Programme umfasst grundsätzlich die folgenden Fragestellungen (Polt et al. 2002; Georghiou, Roessner 2000; Rip 1990): Ermöglichte die evaluierte politische Maßnahme adäquate Lösungen für das zugrunde liegende technische, wissenschaftliche, wirtschaftliche oder gesellschaftliche Problem? Sind die Annahmen des Programms zutreffend? Der Versuch der Beantwortung wirft methodische, konzeptionelle und empirische Probleme auf, die in jedem Fall neu bewältigt werden müssen. Ein entscheidendes methodisches Problem hat seine Ursache in der Praxis der Politikgenerierung: Nur in seltenen Ausnahmefällen werden forschungs- und technologiepolitische Ziele explizit, klar und im Hinblick auf ihre Einlösung kontrollierbar formuliert (Cunningham et al. 1994). Die weiteren Fragestellungen eines Evaluationskonzepts betreffen die Wirkungen und die Durchführung des Programms: Wurde die Zielgruppe erreicht? Welche direkten und indirekten Wirkungen lassen sich dem Programm zuweisen? Wurden die Programmziele "erreicht"? Diese Frage ist keinesfalls trivial, nicht nur wegen der Unklarheit von Zielsetzungen, sondern auch wegen des Problems der Zuschreibung von Sachverhalten zu Anstoßwirkungen des Programms. Neben der Zielerreichung ist außerdem nach "Mitnahmeeffekten" und nach der Angemessenheit der Implementation und der administrativen Abwicklung von Programmen zu fragen.

3.3 Objektivität und Reichweite von Evaluationsverfahren

Auch wenn Evaluationsverfahren heute in weiten Bereichen staatlicher Forschungs- und Innovationspolitik Anwendung finden und akzeptabel sind, so lässt sich doch die "Richtigkeit" und Objektivität ihrer Ergebnisse und der daraus abgeleiteten Empfehlungen mit dem Hinweis auf unzureichende Evaluationsmethodik immer wieder in Zweifel ziehen (z. B. Schön, Rein 1994):

- Die verwendeten Evaluationskriterien vernachlässigen die Tatsache, dass die meisten Programme "multiple, conflicting, and evolving purposes" verfolgen.
- Programmergebnisse werden häufig evaluiert, ohne dass ihr Entstehungskontext hinreichend verstanden worden ist ("black box problem").
- Evaluation erwarb sich vielfach den Ruf eines "Killers", wenn sie den (teilweisen) Misserfolg eines Programms nachwies, ohne sich mit den möglichen Ursachen zu beschäftigen.
- Kritiker von Programmevaluationen verweisen darauf, dass diese die Perspektive der politisch-administrativen Programmverantwortlichen (oder staatlicher Aufsichtsorgane) einnehmen, aber die Interessen sonstiger "Betroffener" außer Acht ließen.

Zu diesen generellen Problemen von Programmevaluationen treten für das Politikfeld "Forschung und Innovation" spezifische hinzu: (1) Forschung und Innovation haben vielfältige unvorhersehbare Effekte. Zu den kurzfristigen Wirkungen für die Teilnehmer geförderter Projekte gehören Wissensgewinn oder Umsatzsteigerungen sowie neuartige Kontakte. Die Effekte von Forschung gehen aber weit darüber hinaus: Ein Projekt übt auch Wirkungen auf Akteure aus, die nicht daran teilgenommen haben - solche Einflüsse sind schwer zu messen. (2) Häufig verlangen politische Entscheidungsträger Evaluationsergebnisse früher als Forschungsarbeiten abgeschlossen sind und sich sozioökonomische Effekte zeigen. (3) Effekte können die Folge einer Kombination von Einflüssen sein (Airaghi et al. 1999). Die offensichtlichen Begrenzungen der Leistungsfähigkeit von Evaluationsverfahren lassen sich nur kompensieren, wenn ihr Verwendungskontext bewusst gehalten und Ergebnisse mit Vorsicht bewertet werden (vgl. Kuhlmann

• Evaluatoren und Evaluationsnutzer müssen sich klar darüber werden, welche *Fragen* in

- welcher Breite, Tiefe und Radikalität zu stellen sind.
- Die Evaluationsforschung im Bereich der Forschungs- und Innovationspolitik geht heute davon aus, dass "realistische" Informationen über ein Evaluationsobjekt nur durch die kombinierte Verwendung verschiedener sozialwissenschaftlicher Methoden und Indikatoren erarbeitet werden können.
- Dennoch sind der "Objektivität" von Evaluationsergebnissen enge Grenzen gesetzt, die durch eine gezielte Berücksichtigung verschiedenartiger Akteursperspektiven allerdings "überwunden" werden können.

Vorsicht und Sorgfalt sind angebracht. Die Innovationsforschung weiß heute, dass die Zusammenhänge zwischen Investition in Forschung und Innovation und möglichen wissenschaftlichen, wirtschaftlichen, sozialen oder ökologischen "Erträgen" komplex sind und keinesfalls als schlichtes Input/Output-Modell konstruiert werden können (z. B. Grupp 1998).

4 Leistungsmessung oder Lernmedium

Die Erwartungen an Evaluationsverfahren bewegen sich dabei zwischen zwei funktionalen Polen: Evaluation kann in erster Linie der *Leistungsmessung* und damit der nachträglichen Rechtfertigung von Fördermaßnahmen dienen (*summative Funktion*), oder sie kann als "*Lernmedium*" verwendet werden, indem sie Erkenntnisse über Ursache-Wirkungszusammenhänge laufender oder abgeschlossener Maßnahmen als intelligente Information für laufende oder künftige Initiativen nutzt (*formative Funktion*).

Der summative Pol ist vor allem in der Evaluationspraxis angloamerikanischer Staaten beheimatet: Hier gewannen im Rahmen der Bemühungen zur Reform und Kostensenkung im öffentlichen Sektor ("New Public Management") auch in der Forschungs- und Innovationspolitik Verfahren der Leistungsmessung ("Performance Measurement") großen Einfluss (Shapira et al. 1997). Angetrieben vom "Government Performance and Results Act (GPRA)" betreiben die US-Regierung und eine Mehrheit der Bundesstaaten zusehends "per-

formance-based management and budgeting systems" (Cozzens 2003; Feller 2003).

Da aber die Komplexität forschungs- und innovationspolitischer Programme wie auch der Aufgaben von Institutionen eher gewachsen ist, stoßen summative Leistungsmessungen schnell an ihre Grenzen. Formative, lernorientierte Evaluationsansätze wurden deshalb – teils in Konkurrenz, teils als Ergänzung zu summativen - ebenfalls weiter entwickelt und eingesetzt. Von der Erfahrung ausgehend, dass Evaluationsergebnisse häufig nur geringe Wirkung in politischen Entscheidungsprozessen zeitigen oder nur wenige der in einer Policy-Arena vertretenen Erwartungshaltungen und Interessenpositionen unterstützen, versuchten Evaluationsexperten (und zunehmend auch policy-maker) die Grenzen zwischen Evaluation und Entscheidungsprozessen zu lockern, ja sogar beide Sphären teilweise zu integrieren. Der Schlüsselbegriff des neuen, erweiterten Evaluationsverständnisses lautet "Verhandlung" in Akteursarenen. Das Ergebnis von Evaluationen, die entsprechend konzipiert wurden, ist, im Unterschied zur konventionellen Methodologie, nicht länger "a set of conclusions, recommendations, or value judgements, but rather an agenda for negotiation of those claims, concerns, and issues that have not been resolved in the hermeneutic dialectic exchanges" (Guba, Lincoln 1989: 13): Hier treten also der Evaluationsprozess und sein medialer Charakter in den Vordergrund; er wird bewusst "partizipativ" gestaltet (Patton 1997; Worthen et al. 1997; Kuhlmann 1998):

- Evaluation wird als Verfahren der empirisch-analytisch aufbereiteten, *strukturierten* Präsentation und Konfrontation von (teilweise widerstreitenden) *Akteursperspektiven* konzipiert; dabei kann das gesamte Spektrum von Evaluationsmethoden zum Einsatz gebracht werden.
- Der Evaluator agiert als "facilitator", er unterstützt die Moderation der Auseinandersetzungen im Verhandlungssystem durch Akteure des politisch-administrativen Systems.
- Das Evaluationsziel ist nicht allein die Bewertung von Sachverhalten aus einer einzelnen Akteursperspektive (z. B. des politischadministrativen Systems), oder die "objektive" Prüfung der Eignung einer Policy, sondern die Stimulation von Lernprozessen

durch Überwindung verfestigter Akteursorientierungen.

Diese Evaluationskonzeptionen zielen vor allem darauf, ein "re-framing" (Schön/Rein 1994) der Orientierungen korporatistischer und politischadministrativer Akteure zu erleichtern. "Intelligente" Politikentwicklungsverfahren in diesem Sinne können darüber hinaus bereichert werden durch Kombination mit (vgl. Kuhlmann 2003b) Vorausschau-Verfahren ("Foresight"; vgl. Cuhls et al. 2002), in der Absicht, diskussionsfähige "Visionen" von mehr oder weniger wünschbaren Zukunftsentwicklungen zu liefern, sowie "Technology Assessment" (vgl. z. B. Rip et al. 1995; Grunwald 2002) als dem Versuch, die möglichen positiven oder negativen Wirkungen technologischer Entwicklungen zu antizipieren und die in solchen Studien gewonnenen Informationen in den Prozess der Technikgenese rückzukoppeln. Dieses Potential methodisch angereicherter akteurs- und prozessorientierter Evaluationsverfahren stößt in der Forschungsund Innovationspolitik in Europa auf wachsendes Interesse.

Anmerkungen

- 1) Der Beitrag ist eine gekürzte und aktualisierte Fassung von Kuhlmann (2000).
- 2) Gibbons u. a. (1994) fassen diese Entwicklung als Übergang von "Mode 1" zu "Mode 2" zusammen. Ähnlich argumentiert auch Mittelstraß 1994. Etzkowitz, Leydesdorff (2000) entwickeln ein "Triple Helix" Modell, nach dem Wissenschaft ko-evolutionär in die Dynamiken des Wirtschaftsund des politischen Systems verflochten ist.
- 3) Die deutsche Evaluationspraxis im Bereich von Forschungs- und Innovationspolitik steht in enger Beziehung zu ähnlichen "Evaluationskulturen" in anderen Industrieländern; Übersichten bei Bozeman, Melkers 1993; Georghiou 1995; OECD 1997; Hirasawa 2002; Shapira, Kuhlmann 2003. Dabei setzte sich in den angelsächsischen Ländern eher eine quantitativ orientierte Leistungsmessung durch, während in Deutschland und Frankreich eher qualitative Verfahren dominieren. In Japan verfolgt man seit Mitte der 1990er offiziell eine Kombination beider Linien, wobei die Implementation noch in den Kinderschuhen steckt.

Zitierte Literatur

Airaghi, A.; Busch, N.E.; Georghiou, L.; Kuhlmann, S.; Ledoux, M.J.; Raan, A.F.J. van; Viana Baptista, J., 1999: Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes. Brussels, Luxembourg: Official Publications of the EC. – ISBN 92-828-3721-1

Block, H.-J.; Krull, W., 1990: What are the consequences? Reflections on the impact of evaluations conducted by a science policy advisory body. In: Scientometrics, Vol. 19, Nos. 5-6, S. 427-437

BMBF (Bundesministerium für Bildung und Forschung), 2002: Facts and Figures Research 2002, Bonn

Bozeman, B; Melkers, J. (eds.), 1993: Evaluating R&D Impacts: Methods and Practice. Boston, Dordrecht, London: Kluwer Academic Publishers

Brockhoff, K., 1999: Forschung und Entwicklung. Planung und Kontrolle. München, Wien: R. Oldenbourg, 5. erg. u. erw. Aufl.

Bürgel, H.D.; Haller, C.; Binder, M., 1996: F&E-Management. München: Vahlen

Callon, M., 1992: The Dynamics of Techno-Economic Networks. In: Coombs, R.; Saviotti, P.; Walsh, V. (eds.): Technological Change and Company Strategies: Economic and sociological perspectives. London et al.: Academic Press Limited, S. 72-102

Cozzens, S., 2003: Frameworks for evaluating S&T Policy in the United States. In: Shapira, P., Kuhlmann, S. (eds.): Learning from Science and Technology Policy Evaluation. Cheltenham: E. Elgar (forthcoming)

Cuhls, K.; Blind, K., Grupp, H., 2002: Innovations for our Future. Delphi '98: New Foresight on Science and Technology. Heidelberg, New York: Physica, Springer

Cunningham, P.N.; Georghiou, L.G.; Barker, K.E.; Kuhlmann, S.; Reger, G.; Marciano da Silva, C.; Henriques, L., 1994: Analysis of Experience in the Use of Verifiable Objectives. Brussels, Luxembourg: Commission of the EC - EUR 15634 EN

Daniel, H.-D., 1993: Guardians of Science. Fairness and Reliability of Peer Review. Weinheim u.a.: VCH DFG (Deutsche Forschungsgemeinschaft), 2002: Reform des Begutachtungssystems; siehe: http://www.dfg.de/organisation/fachgutachter/reform.html

Etzkowitz, H.; Leydesdorff, L., 2000: The dynamics of innovation: from National Systems and "Mode2" to a Triple Helix of university-industry-government relations. In: Research Policy 29, S. 109-123

Evaluierungskommission (Hrsg.), 1998: Systemevaluierung der Fraunhofer-Gesellschaft. Bericht der

Evaluierungskommission. München (http://www.fraunhofer.de/german/company/fhmodel/evaluierungsbericht.pdf)

Feller, I., 2003: The academic policy analyst as reporter: the who, what and how of evaluating science and technology programs. In: Shapira, P.; Kuhlmann, S. (eds.): Learning from Science and Technology Policy Evaluation. Cheltenham: E. Elgar (forthcoming)

Georghiou, L., 1995: Research Evaluation in European National Science and Technology Systems. In: Research Evaluation, Vol. 5, No. 1, S. 3-10

Georghiou, L.; Rigby, J.; Cameron, H. (eds.), 2003: Assessing the Socio-Economic Impacts of the Framework Programme. Manchester, Brussels: PREST/EU Commission, S. 144-180; see also http://www.cordis.lu/fp5/monitoring/studies.htm

Georghiou, L.; Roessner, D., 2000: Evaluating technology programs: tools and methods. In: Research Policy

Gibbons, M.; Limoges, C.; Nowottny, H.; Schwartzman, S.; Scott, P.; Trow, M., 1994: The new production of knowledge. The dynamics of science and research in contemporary societies. London et al.: Sage

Grupp, H., 1998: Foundations of the Economics of Innovation: Theory, Measurement and Practice. Cheltenham: E. Elgar Publishing

Grunwald, A., 2002: Technikfolgenabschätzung – Eine Einführung. Berlin: edition sigma

Guba, E.G.; Lincoln, Y.S., 1989: Fourth Generation Evaluation. Newbury Park et al.: Sage

Guston, D., 2003: The expanding role of peer review processes in the United States. In: Shapira, P.; Kuhlmann, S. (eds.): Learning from Science and Technology Policy Evaluation. Cheltenham: E. Elgar Publishing (forthcoming).

Hirasawa, R., 2002: Framework on Public R&D Evaluation from the Viewpoint of New Public Management: Related to the Present Japanese Situation, International Conference on Socio-economic Evaluation of Public Research and Development - Perspectives, Tokyo, 18./19. April 2002 (http://www.meti.go.jp/policy/tech_evaluation/engli sh/f00/f0000001.html)

Hornbostel, S., 1997: Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag

Internationale Kommission, 1999: Forschungsförderung in Deutschland. Bericht der internationalen Kommission zur Systemevaluation der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft. Hannover: Volkswagen-Stiftung

Kodama, F., 1995: Emerging Patterns of Innovation. Sources of Japan's Technological Edge. Boston: Harvard Business School Press

Kommission Systemevaluation der Wirtschaftsintegrierenden Forschungsförderung, 2001: Endbericht der Kommission. Berlin: mimeo; http://www.bmwi. de/Homepage/download/technologie/ Systemevaluation.pdf

Kuhlmann, S., 1998: Politikmoderation. Evaluationsverfahren in der Forschungs- und Technologie-politik. Baden-Baden: Nomos

Kuhlmann, S., 2000: Evaluation in der Forschungsund Innovationspolitik. In: Stockmann, R. (Hrsg.): Evaluationsforschung. Leverkusen: Leske+Budrich, S. 287-307

Kuhlmann, S., (2003a): Evaluation of Research and Innovation Policies – A discussion of Trends with Examples from Germany. In: International Journal of Technology Management (forthcoming)

Kuhlmann, S., (2003b): Evaluation as a Source of "Strategic Intelligence". In: Shapira, Ph., Kuhlmann, S. (eds.): Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe. Cheltenham: E. Elgar Publishing (forthcoming)

Kuhlmann, S.; Holland, D., 1995a: Evaluation von Technologiepolitik in Deutschland – Konzepte, Anwendung, Perspektiven. Heidelberg: Physica-Verlag

Kuhlmann, S.; Holland, D., 1995b: Erfolgsfaktoren der wirtschaftsnahen Forschung. Heidelberg: Physica-Verlag

Larédo, P.; Mustar, P., 2001: Research and Innovation Policies in the New Global Economy. An International Comparative Analysis. Cheltenham: E. Elgar Publishing

Luhmann, N., 1990: Die Wissenschaft der Gesellschaft. Frankfurt a.M.: Suhrkamp

Lundvall, B.-Å.; Borrás, S., 1998: The globalising learning economy: Implications for innovation policy. Luxembourg: Office for Official Publications of the European Communities

Meyer-Krahmer, F., 1989: Der Einfluß staatlicher Technologiepolitik auf industrielle Innovationen. Baden-Baden: Nomos

Meyer-Krahmer, F., 1999: Was bedeutet Globalisierung für Aufgaben und Handlungsspielräume nationaler Innovationspolitiken? In: Grimmer, K.; Kuhlmann, S.; Meyer-Krahmer, F. (Hrsg.): Innovationspolitik in globalisierten Arenen. Leverkusen: Leske+Budrich, S. 35-65

Meyer-Krahmer, F.; Schmoch, U., 1998: Science-based technologies: university-industry interactions in four fields. In: Research Policy 27, S. 835-851

Mittelstraβ, J., 1994: Grundlagen und Anwendungen – Über das schwierige Verhältnis zwischen Forschung, Entwicklung und Politik. In: Chem.-Ing.Tech. 66, Nr. 3, S. 309-315

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

Noyons, E.; Schmoch, U. et al., 2003: Mapping Excellence in Science and Technology in Europe. Report to the EU Commission, DG Research (forthcoming)

OECD (ed.), 1997: Policy Evaluation in Innovation and Technology, Towards Best Practices. Paris

Patton, M.Q., 1997: Utilization-Focused Evaluation. The New Century Text. Thousand Oaks et al.: Sage

Polt, W.; Rojo, J.; Zinöcker, K.; Fahrenkrog, G.; Tübke, A. (eds.), 2002: RTD Evaluation Tool Box – Assessing the Socio-Economic Impact of RTD-Policies. Brussels, Luxembourg: European Commission, Institute for Prospective Technological Studies (EUR 20382 EN)

Raan, A.F.J. van (ed.), 1988: Handbook of Quantitative Studies of Science and Technology. Amsterdam: Elsevier

Rip, A., 1990: Implementation and Evaluation of Science & Technology Priorities and Programs. In: Cozzens, S. (ed.): The Research System in Transition. Boston, Dordrecht, London: Kluwer, S. 263-280

Rip, A.; Misa, Th.J.; Schot, J. (eds.), 1995: Managing Technology in Society. The Approach of Constructive Technology Assessment. London, New York: Pinter

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

Schön, D.; Rein, M., 1994: Frame Reflection. Toward the Resolution of Intractable Policy Controversies. New York: BasicBooks

Shapira, P.; Kingsley, G.; Youtie, J., 1997: Manufacturing Partnerships: Evaluation in the Context of Government Reform. In: Evaluation and Program Planning, 2, 1, S. 103-112

Shapira, P.; Kuhlmann, S. (eds.), 2003: Learning from Science and Technology Policy Evaluation. Cheltenham: E. Elgar Publishing (forthcoming)

Widmer, T., 1996: Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern et al.: Haupt

Wissenschaftsrat (Hrsg.), 2000: Systemevaluation der Blauen Liste – Stellungnahme des Wissenschaftsrates zum Abschluss der Bewertung der Einrichtungen der Blauen Liste. Köln; http://www.wissenschaftsrat.de/texte/4703-01.pdf

Wissenschaftsrat (Hrsg.), 2001: Systemevaluation der HGF – Stellungnahme des Wissenschaftsrates

zur Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren. Köln;

http://www.wissenschaftsrat.de/texte/4755-01.pdf

Worthen, B.R.; Sanders, J.R.; Fitzpatrick, J.L., 1997: Program Evaluation. Alternative Approaches and Practical Guidelines (2nd ed.). White Plains, N.Y.: Longman

Kontakt

Prof. Dr. Stefan Kuhlmann

Head of Dept. "Technology Analysis & Innovation Strategies"

Fraunhofer Institute for Systems and Innovation Research (ISI)

Breslauer Strasse 48, 76139 Karlsruhe

Tel.: +49 (0) 721 / 68 09 - 170/4 Fax: +49 (0) 721 / 68 09 - 260

E-Mail: s.kuhlmann@isi.fraunhofer.de

Internet: http://www.isi.fhg.de/ti/departm.htm

Auch:

Utrecht University Copernicus Institute Dept. of Innovation Studies Heidelberglaan 8, PO Box 80125 NL-3508-TC Utrecht, The Netherlands

Tel.: +31 - 30 / 253 - 16 25 Internet: http://www.nwi.uu.nl

«