

# Erfolgskontrolle und Lernmedium: Evaluation von Forschungs- und Innovationspolitik

Von Stefan Kuhlmann und Susanne Bühler\*

## Zusammenfassung

*Der vorliegende Beitrag skizziert die Entwicklungen von Evaluationsverfahren in der Forschungs- und Innovationspolitik. Nach einer kurzen Darstellung der Rahmenbedingungen staatlicher Forschungs- und Technologieförderung in Deutschland (Abschnitt 2) erfolgt eine Übersicht der wichtigsten Konzepte, Anwendungsbereiche und Methoden evaluativer Verfahren in diesem Politikfeld (Abschnitt 3), sodann (Abschnitt 4) eine Skizze des anhaltenden strukturellen Wandels in Forschung und Innovation sowie (Abschnitt 5) eine Diskussion seiner Konsequenzen für die Verwendung von Evaluationsverfahren im Prozess politischer Entscheidungsfindung; die Ausführungen werden jeweils durch aktuelle Fallbeispiele illustriert. Ein Ausblick schließt den Beitrag (Abschnitt 6).*

## 1. Einführung

Rund ein Drittel aller Wissenschaft und Forschung in Deutschland wird von der öffentlichen Hand getragen; Unternehmen der Wirtschaft finanzieren die anderen zwei Drittel.<sup>1</sup> Seit den 1990er Jahren erwarten politische Entscheidungsträger, aber auch eine kritische Öffentlichkeit verstärkt Rechenschaft darüber, welcher „Nutzen“ mit dem ersten Drittel, den staatlichen Aufwendungen einhergeht. Grund hierfür ist einerseits die Knappheit öffentlicher Mittel, andererseits aber auch eine gewisse Skepsis gegenüber den Autonomieansprüchen der Wissenschaft, die ihren Ausdruck nicht zuletzt in verschiedenen, seit vielen Jahrzehnten angewendeten Verfahren der internen Selbstevaluation und Selbststeuerung (*peer review*) findet. Große Forschungsprojekte, übergreifende Forschungsförderungsprogramme und Forschungsinstitutionen werden zusehends zum Gegenstand politisch initiiertes, nicht ausschließlich vom Wissenschaftssystem getragener Evaluationsverfahren.<sup>2</sup> Parallel zur Verbreitung und Ausdifferenzierung dieser Versuche der Erfolgsbewertung entfaltete sich auch eine Evaluationsforschung. Sie analysiert die Anwendungsbedingungen sowie Erfolgsvoraussetzungen von Evaluation im Bereich der Forschungs- und Innovationsförderung und entwickelt weiterführende Konzepte und Methoden,<sup>3</sup> neuerdings häufiger auch dahingehend, dass Evaluationsverfahren als „Lernmedium“ für die beteiligten Akteure in

Forschung und Innovation angelegt werden. Der vorliegende Beitrag behandelt beides, die Evaluationspraxis sowie die Evaluationskonzepte in der Forschungs- und Innovationspolitik.

---

\* PD Dr. rer.pol. Stefan Kuhlmann, Fraunhofer-Institut für Systemtechnik und Innovationsforschung (ISI), Karlsruhe, [sk@isi.fhg.de](mailto:sk@isi.fhg.de); Dr. phil. Susanne Bühler, Fraunhofer ISI, [sub@isi.fhg.de](mailto:sub@isi.fhg.de). Der Beitrag enthält überarbeitete und um praktische Beispiele ergänzte Passagen eines Textes (Kuhlmann 2000a), der in einem Band zur Evaluationsforschung veröffentlicht wurde (Stockmann 2000), sowie überarbeitete Teile von Kuhlmann/Bühler (1999).

<sup>1</sup> European Commission (1999).

<sup>2</sup> Auch die Forschungs- und Entwicklungslaboratorien der Industrieforschung werden immer häufiger auf ihre Effizienz, Effektivität und auf ihre strategische Positionierung hin untersucht; auf diese industrielle Evaluationspraxis gehen wir im vorliegenden Beitrag jedoch nicht ein. Darstellungen der entsprechenden Vorgehensweisen finden sich in Brockhoff (1999); Gerpott (1999); Bürgel et al. (1996).

<sup>3</sup> Die Evaluationsforschung profitiert von einer mittlerweile seit Jahrzehnten fortschreitenden, politikfeldübergreifenden Verwissenschaftlichung der Evaluationspraxis. Amerikanische Arbeiten gelten hier als führend (z. B. Guba/Lincoln 1989; Scriven 1991; Patton 1997). Auch im deutschsprachigen Raum gibt es, nach bemerkenswerten Ansätzen in den 1970er und 80er Jahren (Derlien 1976; Hellstern/Wollmann 1984) wieder starke Bestrebungen einer wissenschaftlichen Basierung der Evaluationspraxis (z. B. Widmer 1996; Wollmann 1998; Stockmann 2000).

## 2. Rahmenbedingungen staatlicher Forschungs- und Innovationsförderung in Deutschland

Staatliche Interventionen in Forschung, Technologie und Innovation haben eine lange Tradition und bilden heute unbestritten ein eigenständiges Feld staatlicher Politik.<sup>4</sup> Wenn im Folgenden von Forschungs- und Innovationspolitik die Rede ist, so sind damit alle Anstrengungen des politischen Systems gemeint, das Innovationssystem<sup>5</sup> und seine „Forschungslandschaft“ — die Infrastruktur forschender und Technologie entwickelnder Institutionen einschließlich der ihnen geltenden Regulationen — zu gestalten.<sup>6</sup> Dies schließt, spätestens seit den 1970er Jahren, Bemühungen ein, das Innovationsverhalten industrieller Unternehmen positiv zu beeinflussen.<sup>7</sup> Dabei ist zu betonen, dass Forschungs- und Innovationspolitik wichtige Schnittstellen zu anderen Politikfeldern aufweist, insbesondere zur Bildungspolitik (Aus-, Fort- und Weiterbildung), zur Wirtschaftspolitik (Strukturwandel, Handelspolitik), zur Rechts- und Innenpolitik, zur Umwelt- und Verkehrspolitik. Diese Politikbereiche beeinflussen entweder entscheidende Randbedingungen von Forschung und Innovation auf der Angebotsseite (Infrastruktur, qualifiziertes Personal etc.) oder wirken von der Nachfrageseite her (wie Verkehrsinfrastruktur, Arbeitsschutz, Umweltregulationen).

Das deutsche Forschungs- und Innovationssystem gilt im internationalen Vergleich als relativ gut entwickelt: 1995 waren knapp 460 000 Personen (Vollzeitäquivalente) mit Forschung und Entwicklung (FuE) beschäftigt. Die Gesamtausgaben für FuE beliefen sich auf knapp 81 Mrd. DM; das entspricht 2,35 % des Bruttoinlandsprodukts.<sup>8</sup> Die Forschungsinfrastruktur lässt sich als vergleichsweise differenziert charakterisieren:

- Die *Industrie* realisiert den größten Anteil von Forschung und Entwicklung in Deutschland (283 000 FuE-Mitarbeiter): 1995 investierte die Industrie knapp 50 Mrd. DM überwiegend in angewandte Forschung und experimentelle Entwicklung. Nur einige wenige große multinationale Unternehmen, insbesondere in der Chemie und der elektrotechnischen Industrie, führen selber langfristig anwendungsorientierte Grundlagenforschung durch.
- Auf die *Hochschuleinrichtungen* (335 staatliche bzw. staatlich anerkannte Hochschulen, darunter 113 Universitäten und vergleichbare Einrichtungen; 101 000 FuE-Mitarbeiter) entfällt der zweitgrößte Anteil der Forschungsausgaben (14,4 Mrd. DM). Sie konzentrieren sich auf Grundlagenforschung und auf langfristig anwendungsorientierte Forschung, zum größten Teil finanziert durch die Bundesländer sowie die Deutsche Forschungsgemeinschaft (DFG 1995: 1,8 Mrd. DM). Im Verlauf der 1980er Jahre ist der Anteil industrieller Forschungsaufträge an den Forschungsbudgets einzelner Universitäten, insbesondere technischer, signifikant gewachsen (auf 1,2 Mrd. DM).

- Die 16 „Großforschungseinrichtungen“ der *Helmholtz-Gesellschaft* des Bundes (22 000 FuE-Mitarbeiter) leisten vor allem langfristig orientierte Forschung, die als risikoreich gilt, hohe Kosten verursacht (Anlagen) und große Forschungsteams benötigt. In den vergangenen Jahren haben die Großforschungseinrichtungen ihre Aktivitätsfelder deutlich verändert. Ihr Aktivitätsspektrum reicht heute von der Hochenergiephysik über die Raumfahrttechnologie, die Medizin, die Biotechnologie, die angewandte Mathematik und die Software-Entwicklung bis hin zur Umwelttechnologie. Außerdem betreiben verschiedene Ministerien des Bundes so genannte Ressort-Forschungseinrichtungen zur wissenschaftlich-technischen Unterstützung ihrer Aufgabenerledigung.
- Die Institute der *Max-Planck-Gesellschaft* (9 900 FuE-Mitarbeiter), eine Forschungsorganisation, die auf die im Jahre 1911 gegründete „Kaiser-Wilhelm-Gesellschaft“ zurückgeht, konzentrieren sich auf ausgewählte Felder der Grundlagenforschung in den Natur- und Geisteswissenschaften. Sie beschäftigen sich vor allem mit solchen Forschungsschwerpunkten, bei denen ein großes Erkenntnis- und Entwicklungspotential vermutet wird und die noch nicht in der Universitätsforschung verankert sind, oder die — wegen ihres interdisziplinären Charakters oder des erforderlichen Ressourcenaufwandes — dort keinen Platz finden werden.
- Die Institute der *Fraunhofer-Gesellschaft* (ca. 6 200 FuE-Mitarbeiter) 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 werden. Die Gesellschaft betrachtet sich selbst als „Schnittstelle“ zwischen Wissenschaft und Industrie in Deutschland.

<sup>4</sup> Vgl. hierzu z. B. Grimmer et al. (1992); Mowery (1994); Klodt (1995); Martinsen/Simonis (1995); Kuhlmann (1999).

<sup>5</sup> Ein heute weithin akzeptierter Begriff der Innovationsforschung (Freeman 1987; Lundvall 1992; Nelson 1993; Edquist 1997); das „Innovationssystem“ erstreckt sich danach über Schulen, Universitäten, Forschungsinstitute (Bildungs- und Wissenschaftssystem), industrielle Unternehmen (Wirtschaftssystem), die in diesem Felde tätigen politisch-administrativen und intermediären Instanzen (politischen System) sowie die formellen und informellen Netzwerke der Akteure dieser Institutionen. Als „hybrides System“ repräsentiert es einen Ausschnitt der Gesellschaft, der weit in andere Bereiche hineinstrahlt, etwa über das Bildungswesen, oder über unternehmerische Innovationstätigkeit sowie deren sozio-ökonomischen Effekte (Kuhlmann 1999, 13). Kein Innovationssystem gleicht dem anderen, ebenso wenig wie eine Gesellschaft der anderen.

<sup>6</sup> Hierzu auch Krull/Meyer-Krahmer (1996); Ostry/Nelson (1995); Roobeek (1990).

<sup>7</sup> Vgl. Meyer-Krahmer (1989); Ergas (1987).

<sup>8</sup> Zahlen nach BMBF (1998). Der Spitzenwert lag 1987–89 bei 2,9 %. Zum Vergleich: 1997 erreichte Schweden 3,8 %, Japan 2,9 %, USA 2,7 %.

Sie wurde 1949 gegründet; keine andere Forschungsorganisation in Deutschland ist in den vergangenen 20 Jahren so schnell gewachsen.

- Die Forschungseinrichtungen der *Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF)* führen vor allem angewandte Forschung und experimentelle Entwicklung für sektor-spezifische Bedürfnisse industrieller Unternehmen durch. Ihr Angebot, das teils aus öffentlichen Quellen und teils durch die Industrie finanziert wird, richtet sich insbesondere an kleine und mittlere Unternehmen, die in industrie-sektoralen Forschungsvereinigungen organisiert sind.
- Die Institute der *„Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)“* (ehemals „Blaue Liste“) bilden eine Restkategorie, deren Gemeinsamkeit vor allem darin besteht, dass diese Einrichtungen von Bund und Ländern institutionell gefördert werden (rund 9800 Mitarbeiter). Nach der Verschmelzung der beiden deutschen Forschungssysteme 1990 haben viele ostdeutsche Einrichtungen hier eine förderpolitische Heimat gefunden.

1996 gab der Bund insgesamt 16,7 Mrd. DM für Forschung und Entwicklung aus. Das Spektrum von Instrumenten der staatlichen Forschungs- und Innovationspolitik ist heute weit ausdifferenziert und reicht von der institutionellen Förderung von Forschungseinrichtungen (6,8 Mrd. DM) über verschiedene Formen finanzieller Anreize (Programme; 7,5 Mrd. DM) zur Durchführung von Forschung und experimenteller Entwicklung in öffentlichen oder industriellen Forschungslaboratorien bis zur Gestaltung einer „innovationsorientierten“ Infrastruktur einschließlich der Institutionen und Mechanismen des Technologietransfers (siehe Abbildung 1). Diese Instrumente kennzeichnen die Praxis der Forschungs- und Innovationspolitik in der Bundesrepublik Deutschland seit den 1970er Jahren.

### 3. Evaluationsverfahren

#### 3.1 Konzepte und Anwendungsgebiete

In historischer Perspektive kann man in Deutschland wie auch in anderen industrialisierten Ländern zwei heteroge-

Abbildung 1:

### Instrumente staatlicher Forschungs- und Technologiepolitik

<i>Instrumente im engeren Verständnis</i>	<i>Instrumente im weiteren Verständnis</i>
<p>1. Institutionelle Förderung</p> <ul style="list-style-type: none"> <li>• Großforschungseinrichtungen</li> <li>• Max-Planck-Gesellschaft</li> <li>• Fraunhofer-Gesellschaft</li> <li>• Hochschulen</li> <li>• Andere Einrichtungen</li> </ul>	<p>4. Aus- und Fortbildung</p> <ul style="list-style-type: none"> <li>• Schulen, Hochschulen, Unternehmen</li> </ul>
<p>2. Finanzielle Forschungs- und Innovationsanreize</p> <ul style="list-style-type: none"> <li>• Forschungsprogramme und Verbundprojekte</li> <li>• Innovationsprogramme (Indirekte Förderung)</li> <li>• Risikokapital</li> </ul>	<p>5. „Diskursive“ Maßnahmen</p> <ul style="list-style-type: none"> <li>• Evaluation von Innovationspolitik</li> <li>• Technikfolgenabschätzung</li> <li>• Langfristvisionen</li> <li>• <i>Awareness</i>-Maßnahmen</li> </ul>
<p>3. Sonstige Infrastruktur u. Technologietransfer</p> <ul style="list-style-type: none"> <li>• Information und Beratung für KMU</li> <li>• „Demonstrationszentren“</li> <li>• „Technologiezentren“</li> <li>• Kooperation, Netzwerke</li> </ul>	<p>6. Öffentliche Nachfrage</p>
	<p>7. Benachbarte Politikfelder</p> <ul style="list-style-type: none"> <li>• Industrie- und Wettbewerbspolitik</li> <li>• Regulative Politik, z.B. Beeinflussung der privaten Nachfrage</li> <li>• Sozialpolitik</li> </ul>

Nach: Meyer-Krahmer/Kuntze (1992).

ne Entwicklungslinien von Evaluationsverfahren im Bereich der Forschungs- und Innovationspolitik ausmachen.<sup>9</sup> Sie lassen sich als ein *Drei-Schalenmodell* darstellen:<sup>10</sup>

*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.<sup>11</sup> Peer Review-Verfahren kommen im deutschen Forschungs- und Innovationssystem weithin zur Anwendung, insbesondere bei der *ex ante*-Bewertung von Projekten der grundlagen- und der langfristig anwendungsorientierten Forschung.<sup>12</sup> Dieses Verfahren ist das vorherrschende Evaluationsinstrument der Deutschen Forschungsgemeinschaft (DFG). Die DFG spielt eine zentrale Rolle bei der Förderung von Grundlagenforschung an Universitäten, vorwiegend indem sie einzelnen Forschern Zuwendungen auf Antrag gewährt (so genanntes Normalverfahren). Förderanträge werden von Gutachtern (peers) bewertet, die im Vier-Jahres-Rhythmus von der gesamten wissenschaftlichen Gemeinschaft gewählt werden.<sup>13</sup> Jeder Gutachter ist angehalten, den Antrag allein auf der Grundlage seiner wissenschaftlichen Qualität zu beurteilen.

*Zweite Schale — Programme:* Um diesen Kern der Peer Review-Verfahren herum legte sich eine Schale, die aus Evaluationsstudien forschungs- und innovationspolitischer Programme besteht. Die in Deutschland bisher durchgeführten Studien schließen konzeptionell überwiegend an die Tradition der *Wirkungsforschung* an: Grundlage bilden die vielfältigen und traditionsreichen Arbeiten vor allem der amerikanischen Wirkungsforschung (*impact analysis*), in Verbindung mit Ansätzen der *policy analysis*, die sich zu einem in vielen Politikfeldern (insbesondere der Sozialpolitik) angewendeten Instrument der Politikberatung („policy-analytic movement“<sup>14</sup>) entwickelt hat. Dieses Evaluationskonzept kann als wissenschaftsextern bezeichnet werden, denn es wird überwiegend von Akteuren des politisch-administrativen Systems in Gang gesetzt in der Absicht, die Erreichung politisch gesetzter wissenschaftlicher, technologischer, ökonomischer oder gesellschaftlicher Ziele zu prüfen. Wirkungsanalysen haben sich seit den 1970er Jahren in Deutschland mit der Verbreitung von Programm-Politik in vielen Politikfeldern durchgesetzt.<sup>15</sup> Seit dieser Zeit wurde etwa die Hälfte des Forschungshaushalts in die Projekt- bzw. Programmförderung gelenkt, die andere Hälfte in die institutionelle Förderung. Im Unterschied zur institutionellen Förderung verfolgen Programme gezielt politische Steuerungs- und Gestaltungsabsichten, etwa im Bereich der Förderung von „Schlüsseltechnologien“ oder der Stimulierung von Innovationstätigkeit in der mittelständischen Wirtschaft.<sup>16</sup> Weit reichende Steuerungsansprüche verlangten nach Erfolgskontrolle: Die Programmevaluation und Wirkungsforschung erlebten seither einen nachhaltigen Auf-

schwung.<sup>17</sup> Die Ausbreitung von Evaluationsverfahren ist außerdem eng mit der wachsenden Zahl von strategischen Programmen (der EU-Kommission) zur Förderung von Wissenschaft und Technologie verbunden.<sup>18</sup>

Programmevaluationen können sehr verschiedenartigen Zwecken dienen, abhängig von der jeweiligen Perspektiven der Initiatoren, Betroffenen und Nutzer. Policy-Macher und Ökonomen unterstreichen gerne die Rolle von Evaluationen bei der Rechtfertigung eines staatlichen Förderprogramms: Entsprechende Studien sollen die ökonomischen Effekte einer Maßnahme feststellen — möglichst präzise messbar — und damit zur strategischen Allokation von Ressourcen beitragen. Viele professionelle Evaluatoren, aber auch Policy-Macher, betonen andererseits die „mediale“ Rolle von Evaluationsverfahren, die bereits im Programmverlauf Lerneffekte bei allen Beteiligten erzeugen kann. Idealerweise ergänzen sich diese beiden Orientierungen.<sup>19</sup>

In Deutschland entfaltete insbesondere das Bundesforschungsministerium nennenswerte Evaluationsaktivitäten. Zwischen 1985 und 1993 veranlasste man etwa 50 größere Evaluationsstudien; als Programmevaluatoren treten zumeist unabhängige Forschungsinstitute im Auftrag forschungspolitisch-administrativer Akteure auf. Eine vom Bundesforschungsministerium initiierte „Metaevaluation“ dieser Studien — sie sollte Qualität, Transparenz und Wettbewerb der Evaluationspraxis verbessern — offenbarte eine große Diversität der Programmevaluationen hinsichtlich Methodik, empirischer Grundlegung und thematischer Breite. Die Metaevaluation legte aber auch einen konzeptionellen Kernbestand in der Mehrzahl der Studien frei (siehe auch unten, Abschnitt 3.2), der eine Ausgangsplattform für die Formulierung allgemeiner Mindeststandards bildete,<sup>20</sup> die bis heute allerdings keinen bindenden, sondern nur empfehlenden Charakter haben. Auch in der zweiten Hälfte der 1990er Jahre ließen die

<sup>9</sup> Vgl. zum Folgenden Kuhlmann (2000b), 313/4; Kuhlmann/Bührer (1999), 239–241.

<sup>10</sup> Die hier beschriebene deutsche Evaluationspraxis im Bereich von Forschungs- und Innovationspolitik steht in enger Beziehung zu ähnlichen „Evaluationskulturen“ in anderen Industrieländern; Übersichten geben u. a. Bozeman/Melkers (1993); Georghiou (1995) und (1998); OECD (1997).

<sup>11</sup> Vgl. hierzu Daniel (1993); Hornbostel (1997); van Raan (1988).

<sup>12</sup> Campbell/Felderer (1997).

<sup>13</sup> Vgl. Neidhardt (1988).

<sup>14</sup> Vgl. Schön/Rein (1994), 11.

<sup>15</sup> Vgl. Derlien (1976); Mayntz (1980) und (1983); Hellstern/Wollmann (1984).

<sup>16</sup> Z. B. Kuntze/Hornschild (1995).

<sup>17</sup> Meyer-Krahmer (1989)

<sup>18</sup> Airaghi et al. (1999).

<sup>19</sup> Papaconstantinou/Polt (1997), 10.

<sup>20</sup> Kuhlmann/Holland (1995a).

Evaluationsbemühungen bei der Programmförderung kaum nach. Zusätzlich trat nun auch die Leistungsfähigkeit von Forschungsinstitutionen ins Blickfeld der Policy-Macher (siehe unten).

Im Laufe von mittlerweile 25 Jahren Evaluation forschungs- und innovationspolitischer Programme entstand sogar eine kleine „Evaluations-Szene“ im deutschsprachigen Raum, bestehend aus einer Gruppe von Experten und Instituten auf dem Gebiet der Wirtschafts- und Sozialwissenschaften, die ein relativ breites Spektrum von Evaluationskonzepten, -methoden und -instrumenten anwendet<sup>21</sup> und seit 1998 auch professionell in einer „Deutschen Gesellschaft für Evaluation (DeGEval)“ organisiert ist.

*Dritte Schale — Institutionen:* Hier geht es um die Leistungsfähigkeit ganzer Forschungsinstitutionen. In Deutschland spielen seit langem die Begutachtungen durch den Wissenschaftsrat eine wichtige Rolle; bei der Neustrukturierung der „Forschungslandschaft“ Ostdeutschlands nach der Vereinigung übernahmen sie sogar eine gestaltende Funktion.<sup>22</sup> Seit den 1990er Jahren werden institutionelle Evaluationen immer häufiger durchgeführt.<sup>23</sup> Im Frühjahr 1999 schloss eine internationale Kommission eine „Systemevaluation“ der Deutschen Forschungsgemeinschaft und der Max-Planck-Gesellschaft ab (im Auftrag der Regierungschefs von Bund und Ländern<sup>24</sup>). Zeitlich parallel wurde auch eine Systemevaluation der Fraunhofer-Gesellschaft durchgeführt.<sup>25</sup> Eine Evaluation der Einrichtungen der „Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)“<sup>26</sup> sowie die Systemevaluation der Großforschungseinrichtungen der „Helmholtz-Gesellschaft“ werden im Jahre 2000 abgeschlossen.

Zusammenfassend lässt sich die deutsche Evaluationspraxis im Bereich der Leistungen von Forschung und Forschungsinstitutionen als stark und zersplittert zugleich charakterisieren: stark wegen eines hohen Maßes an Selbstorganisation im Forschungssystem, das Konsens und Verpflichtung zwischen den Forschern fördert; die Bereitschaft zur Mitwirkung an evaluativen Leistungsbeurteilungen ist in den vergangenen Jahren deutlich gestiegen. Die Evaluationspraxis ist aber auch unsystematisch und zersplittert, weil die institutionell orientierten Evaluationsbemühungen politisch, inhaltlich und organisatorisch bisher nur wenig aufeinander abgestimmt werden konnten: Die oben erwähnte „Evaluations-Szene“ hat auf die „governance“ der Forschungs- und Innovationspolitik, welche die Evaluationspraxis letztlich prägt, nur sehr eingeschränkten Einfluss. Es wird sich zeigen, ob die Professionalisierungsbestrebungen der Deutschen Gesellschaft für Evaluation (DeGEval) hier mittelfristig für eine Verbesserung sorgen.

### 3.2 Methoden

Wir kennen heute vielfältige Methoden zur Feststellung erzielter oder erzielter Wirkungen. 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, Längsschnittanalysen, andere statistische Verfahren, Technometrie, Bibliometrie, Peer Review) einzeln oder kombiniert verwendet werden.<sup>27</sup> 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.<sup>28</sup> Der verständliche Wunsch nach einem standardisiert anwendbaren „Indikatoren-Werkzeugkasten“ ist bei der Tendenz zur Verfolgung komplexer politischer Zielsetzungen nicht erfüllbar.<sup>29</sup>

Ein aus der Perspektive der Evaluationsforschung akzeptables, methodisch abgerundetes Konzept einer *ex post*-Evaluation politischer Programme umfasst grundsätzlich die folgenden Fragestellungen:<sup>30</sup> Ermöglichte die evaluierte politische Maßnahme adäquate Lösungen für das zu Grunde liegende technische, wissenschaftliche, wirtschaftliche oder gesellschaftliche Problem, das den Anlass für die politische Intervention gab? Sind die dem Programmkonzept zu Grunde gelegten Annahmen hinsichtlich Problemwahrnehmung und -ursachen zutreffend, ist das Programm in diesem Sinne also „strategisch effizient“? Sind mögliche, im Förderfeld beobachtbare Strukturveränderungen Effekte der Maßnahme oder haben solche Strukturveränderungen ihrerseits die potentiellen Politikwirkungen beeinflusst? Der Versuch der Beantwortung wirft eine Reihe von methodischen, konzeptionellen und empirischen Problemen auf, die in jedem Fall neu bewältigt werden müssen. In den meisten Fällen erweist es sich, dass sie nur innerhalb wohl definierter konzeptioneller Grenzen überhaupt beherrschbar sind.

<sup>21</sup> Vgl. Becher/Kuhlmann (1995).

<sup>22</sup> Block/Krull (1990).

<sup>23</sup> Kuhlmann/Holland (1995b).

<sup>24</sup> Vgl. Bund-Länder-Kommission (1999).

<sup>25</sup> Vgl. Fraunhofer-Gesellschaft (1998).

<sup>26</sup> Röbbcke/Simon (1999).

<sup>27</sup> Ausführlich hierzu z. B. Meyer-Krahmer (1989), 60–71; Beiträge in Bozeman/Melkers (1993); Grupp/Kuntze/Schmoch (1995); Grupp (1998).

<sup>28</sup> Vgl. beispielsweise Alewell (1988) zur Evaluation von Hochschuleinrichtungen.

<sup>29</sup> Airaghi et al. (1999).

<sup>30</sup> Hierzu ausführlich Kuhlmann (1998a); Dreher (1997); Kuhlmann/Holland (1995a); Callon et al. (1995); Meyer-Krahmer (1989), 50–60; Rip (1990).

Ein entscheidendes methodisches Problem hat seine Ursache unmittelbar in der Praxis der Politikgenerierung: Nur in seltenen Ausnahmefällen werden forschungs- und technologiepolitische Ziele explizit, klar und im Hinblick auf ihre Einlösung kontrollierbar formuliert.<sup>31</sup>

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 zusätzlich wegen des Problems der Zuschreibung von feststellbaren Sachverhalten zu Anstoßwirkungen des Programms. Neben der Zielerreichung ist außerdem nach „Mitnehmereffekten“ und nach der Angemessenheit der Implementation und der administrativen Abwicklung von Programmen zu fragen. Diese Elemente betreffen die operative Effizienz der evaluierten Maßnahme und gehören zu den methodisch am weitesten entwickelten der Evaluationspraxis.

Auch wenn Evaluationsverfahren heute in weiten Bereichen staatlicher Forschungs- und Innovationspolitik Anwendung finden und Akzeptanz genießen, 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. Bezweifelt wird die methodische Seriosität von Evaluationsstudien sowohl von Policy-Machern als auch von Sozialwissenschaftlern. Schön/Rein (1994)<sup>32</sup> identifizieren die folgenden Probleme:

- 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 vielfach den Ruf eines „Killers“, wenn sie den (teilweisen) Misserfolg eines Programmes nachwies, ohne sich mit den möglichen Ursachen zu beschäftigen.
- Viele Kritiker von Programmevaluationen verweisen darauf, dass diese in der Regel die Perspektive der politisch-administrativen Programmverantwortlichen (oder staatlicher Aufsichtsorgane) einnehmen, aber die Interessen sonstiger „Betroffener“ (z. B. Bürgergruppen) außer Acht lassen.

Zu diesen generellen Problemen von Programmevaluationen treten für das Politikfeld spezifische:<sup>33</sup> (1) Forschung und Innovation haben vielfältige Effekte. Zu den typischen kurzfristigen wünschbaren Wirkungen für die Teilnehmer geförderter Projekte gehören Umsatzsteigerungen und vergrößerte Marktanteile, Verbesserungen des Know-how sowie neuartige Kontakte. Die Effekte von

Forschung gehen aber weit darüber hinaus: Ein Projekt übt auch vielfältige Wirkungen auf Akteure aus, die nicht unmittelbar daran teilgenommen haben. Solche Einflüsse sind allerdings schwer zu messen. (2) Evaluationsergebnisse werden häufig frühzeitig für politische Entscheidungen benötigt, manchmal sogar bevor die Forschungsarbeiten abgeschlossen sind und normalerweise bevor sich die gesamte Breite möglicher sozio-ökonomischer Effekte entfaltet haben kann. (3) Viele Effekte lassen sich nicht eindeutig einem bestimmten Projekt oder Programm zuordnen. Sie können die Folge einer Kombination von Einflüssen sein, wozu auch die Verfahren der praktischen Umsetzung von Forschungsergebnissen zählen.

Ein Beispiel, wie mit den oben genannten Schwierigkeiten umzugehen versucht wird, ist die *kontinuierliche und systematische Evaluation (CSE)* der EUREKA-Initiative.<sup>34</sup> Als Konsequenz dreier umfassender Evaluationen Anfang der 1990er Jahre<sup>35</sup> wurde im Jahre 1997 ein „Evaluations-system“ implementiert, das auf folgenden Elementen basiert:

- einer kontinuierlichen Befragung aller Teilnehmer abgeschlossener Projekte mit Hilfe eines schriftlichen standardisierten Fragebogens,
- halbstrukturierten Interviews mit Projektleitern abgeschlossener Projekte, und
- einer Bewertung der Ergebnisse aus den Befragungen und Interviews durch eine international zusammengesetzte Expert Advisory Group.<sup>36</sup>

Die Besonderheit des Vorgehens liegt darin, dass die CSE nicht nur direkt nach Beendigung der Projekte Informationen einholt (klassische ex post-Evaluation), sondern in regelmäßigen Intervallen weitere Erhebungen durchführt. Hierzu werden zwei Arten von Fragebögen verwendet:

<sup>31</sup> Ausführlich untersuchten diesen Zusammenhang in verschiedenen Industrieländern Cunningham et al. (1994); siehe auch Mowery (1994), 10.

<sup>32</sup> Schön/Rein (1994), 12–13.

<sup>33</sup> Airaghi et al. (1999).

<sup>34</sup> EUREKA wurde 1985 als Initiative zur Förderung anwendungsnahe Forschung in Europa gegründet und stellt einen Rahmen zur Förderung der grenzüberschreitenden europäischen Zusammenarbeit von Unternehmen, Forschungseinrichtungen und Hochschulen auf dem Feld fortgeschrittener Technologien bereit. Das Hauptziel besteht in der Förderung der europäischen Wettbewerbsfähigkeit. EUREKA-Projekte zielen darauf, Produkte, Verfahren oder Dienstleistungen zu entwickeln, die von den Projektpartnern gewinnbringend vermarktet werden können und damit die Wettbewerbsfähigkeit der Beteiligten erhöhen. Die Besonderheiten der Initiative sind der „bottom-up“-Ansatz, die Themenvielfalt und ein vergleichsweise geringer administrativer Aufwand. Die Finanzierung der Projekte erfolgt durch die Teilnehmer selbst bzw. durch öffentliche (nationale) Förderstellen.

<sup>35</sup> Vgl. Dekker et al. (1991); Ormala et al. (1993); Airaghi et al. (1995).

<sup>36</sup> Vgl. Bobe et al. (1999).

- *Final Reports (FR)*: Diese werden direkt nach Beendigung eines Projektes an die Teilnehmer versandt und enthalten Fragen zur Struktur des Konsortiums, angestrebten und erreichten Zielen, Finanzierung, industrieller und kommerzieller Verwertung der Ergebnisse, Auswirkungen auf Beschäftigung, wahrgenommener Nutzen der Initiative etc.
- *Market Impact Reports (MIR)*: Diejenigen Befragten, die angeben, eine kommerzielle Verwertung der Projektergebnisse anzustreben, erhalten ein (MIR1), drei (MIR3) sowie fünf Jahre (MIR5) nach Beendigung des Projektes einen kurzen Fragebogen, der nach den tatsächlichen kommerziellen bzw. Beschäftigungsbezogenen Wirkungen fragt sowie Informationen über Ereignisse nach der FuE-Phase (z. B. geänderte Marktbedingungen, Veränderungen in der Unternehmensstrategie etc.) erhebt.

Mit dieser Vorgehensweise ist es möglich, Effekte festzustellen, die längere Zeit bis zu ihrer Realisierung in Anspruch nahmen. Im Rahmen der halbstrukturierten Interviews können dann gezielt Erfolgsfaktoren, aber auch Hindernisse für gelungene kommerzielle Anwendungen eruiert werden.

Insgesamt betrachtet lassen sich die offensichtlichen Begrenzungen der Leistungsfähigkeit von Evaluationsverfahren nur kompensieren, wenn ihr Verwendungskontext bewusst gehalten und ihre Ergebnisse mit Vorsicht bewertet werden. Eine weitere Möglichkeit, die Begrenzungen der Objektivität von Evaluationsergebnissen zu überwinden, besteht in der gezielten Berücksichtigung verschiedenartiger Akteurperspektiven.<sup>37</sup> Für einen „Multi-Akteur-Ansatz“ in der Evaluation staatlicher Maßnahmen sprechen anhaltende, tief greifende Wandlungsprozesse im Innovationssystem.

#### 4. Innovationssysteme im Wandel

Moderne Innovationssysteme standen immer und stehen gegenwärtig ganz besonders unter dem Druck der Anpassung an veränderte Umwelten:<sup>38</sup> Globalisierte Märkte, Umwälzungen der geopolitischen Situation sowie weltweit verflochtene, neuartige Kommunikationsinfrastrukturen und Technologien verändern das Innovationsgeschehen, das wiederum seinerseits durch eigendynamischen Wandel den internationalen Verflechtungsprozess beeinflusst:

- Die Entwicklung von Hochtechnologieprodukten hat in den vergangenen Jahren signifikante Beschleunigung erfahren; die Hervorbringung einer neuen Idee und ihre Kommerzialisierung am Markt folgen heute so schnell aufeinander wie nie. Veränderte Markterfordernisse und der Einsatz von Informations- und Kommunikationstechnik haben wesentlich dazu beigetragen.

- In vielen Gebieten wächst die „Wissenschaftsbindung“ des Innovationsgeschehens; die Funktionen und Verfahrensweisen der „zielgerichteten“ Grundlagenforschung und ihre Schnittstellen zur Entwicklung von marktfähigen Produkten erfordern wachsende Aufmerksamkeit.<sup>39</sup> Ein oft verwendetes Beispiel hierfür bildet die neue Bedeutung molekularbiologischer Forschungen für die pharmazeutische Industrie und die praktizierende Medizin.
- Dabei erfordern komplexe Produkte und Prozesse eine dichtere Verflechtung und „Fusion“<sup>40</sup> heterogener Technikentwicklungen, wodurch traditionelle Grenzziehungen zwischen Wissens- und Technikgebieten verschwinden und interdisziplinäre Kompetenzen im Innovationsprozess an Bedeutung gewinnen,<sup>41</sup> neue disziplinäre Cluster entstehen und alte verschwinden.
- Die genannten Trends steigern insgesamt das Erfordernis der Kooperation von Innovationsakteuren: Mit wachsender Komplexität des erforderlichen Wissens sind isolierte Akteure immer weniger in der Lage, dieses ohne externe Unterstützung hinreichend zu beherrschen.<sup>42</sup> Innovationsorientierte Kooperation und die Pflege entsprechender Netzwerke<sup>43</sup> gehören mittlerweile zum alltäglichen Innovationsgeschehen — zwischen Unternehmen sowie im Rahmen gemeinsamer Forschungs- und Entwicklungsvorhaben zwischen Unternehmen und öffentlichen Forschungseinrichtungen und zusehends auch in internationalen Verflechtungen.<sup>44</sup>
- Nur scheinbar im Widerspruch hierzu steht ein Trend zur Herausbildung von besonders starken, an einen bestimmten Standort gebundenen, thematisch fokussierten Innovationskapazitäten, die weltweit im Wettbewerb mit nur ein oder zwei anderen Standorten stehen, wobei

<sup>37</sup> Vgl. Kuhlmann (1998a), 97–106.

<sup>38</sup> Zum Folgenden Kuhlmann (1999), 19–25.

<sup>39</sup> Schmoch et al. (1996a); Meyer-Krahmer/ Schmoch (1998).

<sup>40</sup> Kodama (1995).

<sup>41</sup> Schmoch et al. (1996b).

<sup>42</sup> Hierzu auch Grupp/Schmoch (1992); van der Meulen/Rip (1994); Felt u. a. (1995), 211–218; Rammert (1997); Kuhlmann/Reger (1996), 74–78; Meyer-Krahmer (1997). Gibbons u. a. (1994) fassen diese Entwicklung als Übergang von „Mode 1“ zu „Mode 2“ zusammen: In „Mode 1 problems are set and solved in a context governed by the largely academic interests of a specific community. By contrast, Mode 2 knowledge is carried out in a context of application. Mode 1 is disciplinary while Mode 2 is transdisciplinary. Mode 1 is characterised by homogeneity, Mode 2 by heterogeneity. Organisationally, Mode 1 is hierarchical and tends to preserve its form, while Mode 2 is more heterarchical and transient“ (Gibbons u. a. 1994, 3). Etzkowitz/ Leydesdorff (2000) entwickeln ein „Triple Helix“ Modell, nach dem Wissenschaft ko-evolutionär in die Dynamiken des Wirtschafts- und des politischen Systems verflochten ist.

<sup>43</sup> Callon (1992).

<sup>44</sup> Jungmittag et al. (1999); Niosi (1999); Gerybadze et al. (1997).

sich die Rangfolge durchaus ändern kann (z. B. Automobilbau in Deutschland und den USA; gentechnologische Pharmazeutik in den USA; Heimelektronik und Bürotechnik in Japan; Segmente der Werkzeugmaschinenindustrie in Italien usw.). Man hat diese „locations“ anschaulich auch als „Industrial Hollywoods“ beschrieben:<sup>45</sup> Wer erfolgreich innovieren will, muss sich an den Maßstäben dieser Zentren, die ein besonders leistungsfähiger Teil regionaler, nationaler oder sektoraler Innovationssysteme sind, orientieren und gegebenenfalls mit ihnen kooperieren.

- Einer neuen Dynamik ist schließlich auch das Verhältnis von kodifiziertem und nicht-kodifiziertem Wissen im Innovationsprozess ausgesetzt; diese Unterscheidung bezieht sich auf das Ausmaß, in dem Wissensbestände fixiert und transferiert werden können. 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.<sup>46</sup>

Staatliche Politik reagierte 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 Programme der Forschungs- und Innovationspolitik als thematisch ausgerichtete Wettbewerbe ausgeschrieben, die auch einen strukturellen Wandel in Wissenschaft und Wirtschaft bewirken sollen: Konsortien von Bewerbern (i. d. R. Institutionen) sollen in einem selbst organisierten Prozess gemeinsame Projekte und deren Detailziele formulieren. Partnerschaften verschiedenster Einrichtungen (Forschungseinrichtungen, Universitäten, private Firmen, Technologiezentren, Weiterbildungseinrichtungen etc.) zielen mit einem ganzen Bündel an aufeinander abgestimmten Maßnahmen auf eine Effektivierung ganzer Innovationssysteme. Solche Multi-Akteur- und/oder Multi-Maßnahmen-Programme können sich entweder auf Regionen und/oder bestimmte Missionen und/oder bestimmte Technologien und Branchen beziehen.
- Die veränderten Anforderungen an die Einrichtungen und Akteure des Innovationssystem schlugen sich auch in den erwähnten „Systemevaluationen“ ganzer Institutionen nieder: So forderten etwa die Gutachter der Systemevaluation der Max-Planck-Gesellschaft und der Deutschen Forschungsgemeinschaft unter anderem eine Lockerung der starken 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 außer-

universitären Forschungseinrichtungen, die Förderung von Institutionen übergreifenden Forschungszentren sowie eine verstärkte internationale Orientierung und Vernetzung der Einrichtungen.<sup>47</sup>

Ein aktuelles Beispiel für Multi-Akteur- und/oder Multi-Maßnahmen-Programme bildet die Fördermaßnahme des Bundesministeriums für Bildung und Forschung „Kompetenzzentren der Nanotechnologie“, in deren Rahmen seit 1998 sechs in einem Wettbewerb ausgewählte Kompetenzzentren (CC) gefördert werden. Die Förderung besteht in erster Linie in der Finanzierung der Infrastruktur der sechs Kompetenzzentren. Darüber hinaus stehen Mittel für die Projektförderung zur Verfügung, die jedoch im Rahmen öffentlicher Ausschreibungen erst eingeworben werden müssen. Die institutionelle Förderung ist zunächst auf drei Jahre begrenzt. In Abhängigkeit der Ergebnisse der begleitenden Evaluation behält sich der Förderer vor, die CC bzw. einige von diesen weitere zwei Jahre zu fördern.

Anlass zur Fördermaßnahme gab die Erkenntnis, dass es für die umfassende und erfolgreiche Nutzung des Potentials der Nanotechnologie notwendig ist, *interdisziplinäre, intersektorale und interinstitutionelle Zusammenarbeit in Netzwerken* unterschiedlicher Akteure aus Wissenschaft, angewandter Forschung und Industrie zu fördern. Dies gilt insbesondere vor dem Hintergrund, dass die Bundesrepublik auf dem Gebiet der Nanotechnologie derzeit (noch) auf einem zweiten Platz, hinter den USA aber vor Japan, angesiedelt wird.<sup>48</sup> Diese Stärken zu nutzen und auszubauen, ist das übergeordnete strategische Ziel der Fördermaßnahme: Bestmögliche Umsetzung von (nanotechnologischem) Wissen in marktfähige Produkte, Verfahren und Dienstleistungen. Darüber hinaus will das CC-Konzept vorhandene Kompetenzen so bündeln, dass der (Nanotechnologie-) Standort Bundesrepublik international sichtbar sowie für nationale und internationale Unternehmen attraktiv gemacht wird. Schließlich sollen die geförderten Kompetenzzentren Aktivitäten auf folgenden Gebieten entfalten: Aus- und Weiterbildung, Öffentlichkeitsarbeit, Standardisierung, Metrik und Normung.

Auf Grund der nicht unerheblichen Summe öffentlicher Mittel für die Fördermaßnahme, der Neuartigkeit des Förderinstruments, der Vielschichtigkeit der Förderziele und -instrumente sowie nicht zuletzt wegen des erkanteten Scheiternsrisikos hat das BMBF eine begleitende Evaluation der Kompetenzzentren der Nanotechnologie veranlasst, die zwei Funktionen erfüllen soll: Sie soll einerseits dabei helfen, Erfahrungen mit dem neuen Förderkonzept zu sammeln, um es auf dieser Basis weiter-

<sup>45</sup> Meyer-Krahmer (1999).

<sup>46</sup> Lundvall/Borrás (1998), 31.

<sup>47</sup> Bund-Länder-Kommission 1999.

<sup>48</sup> Vgl. Zukunft nachgefragt — Juni 1998.



entwickeln zu können; andererseits soll sie die Wirkungen, die das Förderinstrument entfaltet sowie das Ausmaß der Zielerreichung drei Jahre nach Beginn der Förderung feststellen. Damit kommt den Evaluatoren eine Doppelfunktion als „critical friend“ und „objective assessor“ zu. Als Leitprinzip der begleitenden Evaluation gilt, dass keine starre Output-Messung anhand vorgegebener Leistungsparameter vorgenommen wird, sondern die typischen Ausgangs- und institutionellen, technologischen und ökonomischen Rahmenbedingungen der CC einbezogen und berücksichtigt werden. Die Bewertung des Erfolges der Kompetenzzentren ist im Wesentlichen bestimmt durch den Vergleich zwischen Ausgangslage und Veränderungen, also zwischen Statusanalyse zu Beginn und ex post-Evaluation am Ende der Förderzeit. Um zu einer angemessenen Einschätzung der Stärken, Schwächen sowie hemmender und fördernder Faktoren der Entwicklung der CC zu gelangen, sind zusätzliche Mittel der Erfolgsbewertung ein Vergleich mit einem nicht-geförderten Kompetenzzentrum sowie ein internationaler Vergleich.

### 5. Leistungsmessung oder Lernmedium: Evaluation zwischen zwei funktionalen Polen

Die skizzierten Veränderungen der Funktionsbedingungen von Forschung und Innovation haben seit den 1990er Jahren zu einem wachsenden Interesse an Evaluation geführt und der Anwendung entsprechender Verfahren Auftrieb gegeben.<sup>49</sup> 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 wird vor allem von der Evaluationspraxis angloamerikanischer Staaten genährt: 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.<sup>50</sup> Die US-Regierung und eine Mehrheit der Bundesstaaten betreiben zusehends „performance-based management and budgeting systems“ — nicht zuletzt in der Forschungs- und Innovationsförderung.<sup>51</sup>

Da aber die Komplexität forschungs- und innovationspolitischer Programme wie auch der Aufgaben von Institutionen eher gewachsen als zurück gegangen ist, stoßen summative Leistungsmessungen schnell an ihre Gren-

zen. Formative, lernorientierte Evaluationsansätze wurden deshalb — teils in Konkurrenz, teils als Ergänzung zu summativen — ebenfalls weiter entwickelt und eingesetzt. Von der nachhaltigen Erfahrung ausgehend, dass Evaluationsergebnisse häufig nur geringe Wirkung in politischen Entscheidungsprozessen zeitigten und nur wenige der in einer Policy-Arena vertretenen Erwartungshaltungen und Interessenpositionen unterstützten, versuchten Evaluationsexperten (und zunehmend auch Policy-Macher) 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“<sup>52</sup>: Entscheidungen erfolgen eher als fortlaufender Prozess, in welchem konkurrierende Akteure interaktiv Konsens erzielen, oder auch nicht. Evaluationsergebnisse sind dabei eine Information unter vielen. Hier treten also der Evaluationsprozess, genauer die Kommunikationen der beteiligten Akteure in seinem Verlauf in den Vordergrund; der Prozess wird bewusst „partizipativ“ gestaltet („Participatory Evaluation“<sup>53</sup>). Der *mediale* Charakter des Evaluationsverfahrens tritt in den Vordergrund. Insbesondere folgende Eigenschaften des partizipativen Ansatzes lassen sich für den Einsatz in forschungs- und innovationspolitischen Auseinandersetzungen weiterentwickeln:

- Evaluation wird als Verfahren der empirisch-analytisch aufbereiteten, strukturierten Präsentation und Konfrontation von (teilweise widerstreitenden) Akteurperspektiven konzipiert; dabei kann das gesamte Spektrum von Evaluationsmethoden (siehe Abschnitt 3) zum Einsatz gebracht werden.
- Der Evaluator agiert als „facilitator“, er unterstützt die Moderation der Auseinandersetzungen im Verhandlungssystem durch Akteure des politisch-administrativen Systems (siehe Abbildung 2).
- Das Evaluationsziel ist nicht allein die Bewertung von Sachverhalten aus einer einzelnen Akteurperspektive (z. B. des politisch-administrativen Systems), oder die „objektive“ Prüfung der Eignung einer Policy, sondern

<sup>49</sup> Vgl. als Übersicht: OECD (1997).

<sup>50</sup> Osborne/Gaebler (1993); Shapira et al. (1997).

<sup>51</sup> Cozzens/Melkers (1997).

<sup>52</sup> Guba/Lincoln (1989), 13.

<sup>53</sup> Siehe Patton (1997), 100; Worthen et al. (1997), 153–170.

die Stimulation von Lernprozessen durch Überwindung verfestigter Akteursorientierungen.

Die begleitende Evaluation von „Interdisziplinären Zentren für klinische Forschung an Hochschulkliniken“<sup>54</sup> hat einen solchen „medialen“ Charakter. Seit 1995 fördert das Bundesministerium für Bildung und Forschung (BMBF) acht Modellzentren für interdisziplinäre klinische Forschung an den Hochschulkliniken (im Folgenden: IZKF). Die Klinischen Forschungszentren wurden zu dem Zweck gegründet, die lange Zeit wenig befriedigende Qualität klinischer Forschung in Deutschland durch die Entwicklung neuer organisatorischer Strukturen in der Hochschullandschaft zu verbessern.<sup>55</sup> Eine nachhaltige Verankerung dieser Strukturen ist dabei von vier Faktoren abhängig: (1) von der institutionellen Einbettung in den Hochschulen und Hochschulkliniken, (2) von den zur Verfügung stehenden Ressourcen, (3) von der Erwartungssicherheit der am Zentrum arbeitenden Akteure in Bezug auf Forschungsmöglichkeiten und (4) von der Motivation der Teilnehmer, die Ziele der Fördermaßnahme aktiv mitzutragen.

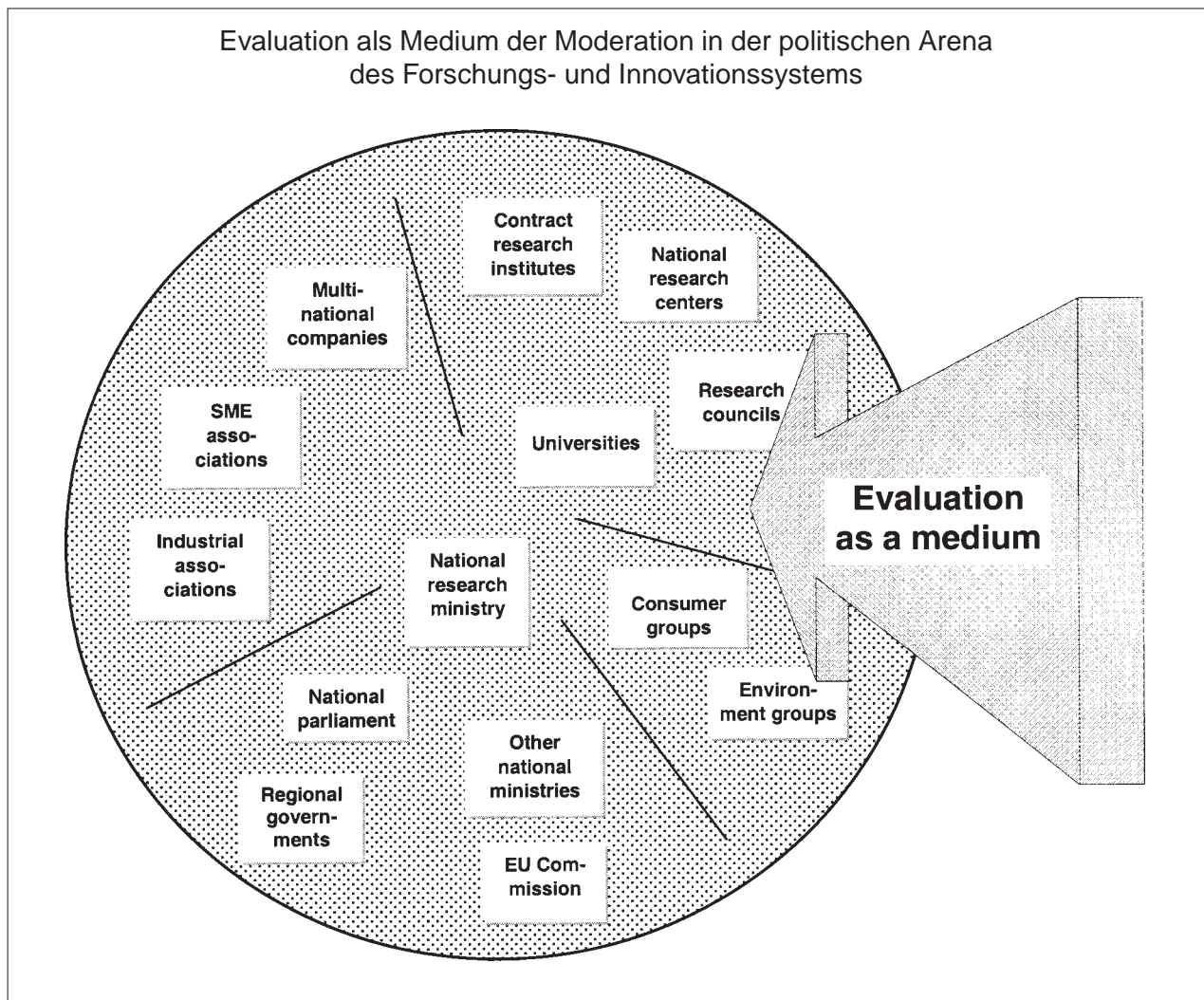
Im Rahmen der begleitenden Evaluation wurden verschiedene Aspekte des angestrebten strukturellen Wandels in der klinischen Forschung untersucht. Unter anderem arbeiteten die Evaluatoren mit Hilfe einer Analyse der Kommunikations- und Kooperationsstrukturen innerhalb der IZKF<sup>56</sup> heraus, ob die Teilnehmer am Zentrum dieses tatsächlich als eine eigene Organisation wahrnehmen, wie sie sich integriert fühlen und mit dem Zentrum identifizieren können, ob die interdisziplinäre Orientierung und allgemeine Kommunikation über organisatorische Maßnahmen angeregt wurde oder nicht, und ob der Nachwuchs etwa durch die Übergabe von Verantwortung und flache Hierarchien in seiner Selbständigkeit und damit

<sup>54</sup> Vgl. Braun et al. (1997).

<sup>55</sup> Mit der Fördermaßnahme sind im Einzelnen folgende Ziele verbunden: (1) Aufbau effizienter Forschungsstrukturen, (2) Steigerung der Interdisziplinarität und der wissenschaftlichen Qualität, (3) Verbesserung der Nachwuchsförderung, (4) Transparenz der Finanzierungsstrukturen.

<sup>56</sup> Vgl. Bühner et al. (1999).

Abbildung 2



Motivation gefördert wurde. Ein zentrales Ergebnis dieser Untersuchung lautete, dass die befragten Kliniker und Forscher, je höher sie auf der Hierarchieebene angesiedelt waren, um so eher dazu neigten, ihre „sozialen Kreise“ innerhalb der eigenen Gruppe zu ziehen bzw. vor allem diese als Kooperationspartner wahrzunehmen. Die Analyse der fächerübergreifenden *Zusammenarbeit* zeigte, dass die professionellen Netzwerke der Kliniker zu jeweils über zwei Dritteln aus Angehörigen klinischer Forschungsorientierungen bestanden. Die Netzwerke der befragten Naturwissenschaftler wie auch der Vorkliniker hingegen waren wesentlich interdisziplinärer angelegt, auch wenn dort ebenfalls mehrheitlich mit Angehörigen derselben Forschungsorientierung kooperiert wird. Diese Ergebnisse widersprachen deutlich den Zielen der Fördermaßnahme und sorgten erwartungsgemäß für intensive Diskussionen innerhalb der IZKF, aber auch mit dem Förderer. Darüber hinaus konnten aber auch zahlreiche Beispiele für gelungene Kommunikation und Kooperation in den Zentren aufgezeigt werden, die als *best practices* einer Reflektion über Optimierungsmöglichkeiten dienten. Mit anderen Worten: Die Kommunikations- und Kooperationsanalysen gab den IZKF selbst, aber auch dem Förderer Informationen an die Hand, die für die Beteiligten durchaus kontraintuitiv waren und dafür gesorgt haben, sich intensiv mit den Zielen der Fördermaßnahme und Strategien der Zielerreichung auseinander zu setzen.

Derartige Evaluationskonzeptionen zielen vor allem darauf, ein „re-framing“<sup>57</sup> der Orientierungen korporatistischer und politisch-administrativer Akteure zu erleichtern. Im Kontext des Forschungs- und Innovationssystems können sie ein „intelligenter“ Zulieferer zu den Verhandlungs- und Bewältigungsstrategien der verantwortlichen politischen Akteure wie der interessierten Öffentlichkeit sein. „Intelligente“ Politikentwicklungsverfahren in diesem Sinne<sup>58</sup> können darüber hinaus bereichert werden durch Kombination mit<sup>59</sup>

- „Vorausschau-Verfahren“ („Technology Foresight“<sup>60</sup>), mit der Absicht, diskussionsfähige „Visionen“ von mehr oder weniger wünschbaren Zukunftsentwicklungen zu liefern, und
- „Technology Assessment“<sup>61</sup> 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 akteur- und prozessorientierter Evaluationsverfahren stößt in der Forschungs-, Technologie- und Innovationspolitik in Europa auf wachsendes Interesse. Aktuelle Beispiele für solche Konzepte finden sich bei den oben dargestellten Maßnahmen des Bundesforschungsministeriums zur strukturellen Modernisierung der interdisziplinären Forschung an deutschen Universitätskliniken, bei der Förde-

rung von komplexen „Kompetenzzentren der Nanotechnologie“, aber auch bei der Maßnahme „Existenzgründer aus Hochschulen“, die Initiativen einer engen regionalen Zusammenarbeit von Wissenschaft und Wirtschaft, von Hochschulen und deren Kooperationspartnern in der Region (wie z. B. Technologie- und Gründerzentren, Unternehmen, Kreditinstitute, Kammern, außeruniversitäre Forschungseinrichtungen) unterstützt, die zum Ziel haben, innovative Unternehmensgründungen von Studierenden sowie Hochschulabsolventen zu fördern und zu betreuen.

## 6. Ausblick

Forschung und Innovation gelten als entscheidende Antriebskräfte der wirtschaftlichen Entwicklung und des Wohlstands der Gesellschaft. Die staatliche Förderung von Forschung und Innovation steht allerdings — abgesehen von der Grundlagenforschung — immer wieder unter Rechtfertigungsdruck, denn sie liegt mit anderen Politikthemen im Wettbewerb um knappe Steuergelder; außerdem sind Objekte, Ziele und Nutzen mancher Fördermaßnahme umstritten. Vor diesem Hintergrund gewann die Evaluation als Bewertung und Erfolgskontrolle staatlicher Forschungs- und Innovationspolitik seit den 1970er Jahren zusehends an Bedeutung.

Jenseits des innerwissenschaftlichen Peer Review-Verfahrens zur Bewertung der wissenschaftlichen Leistungen von Fachkollegen erfolgt Evaluation im Bereich der Forschungs- und Innovationsförderung überwiegend als kritisch prüfender Eingriff von außen: Policy-Macher veranlassen Bewertungen des Erfolges von Fördermaßnahmen gegenüber Wissenschaft und Innovationssystem. Solche Erfolgskontrolle kann als Frage nach dem „return on investment“ von Steuergeldern konzipiert werden — dann droht die Gefahr, kurzfristig die Vielfalt möglicher langfristiger und indirekter, wünschbarer aber auch abzulehnender Wirkungen kreativer Wissenschaft und Innovationstätigkeit in Wirtschaft und Gesellschaft zu übersehen.

Dabei ist das Innovationsgeschehen (und unser Verständnis davon) in den vergangenen Jahren deutlich komplexer geworden: Innovationsprozesse erfolgen als Wechselspiel und Lernprozess einer Vielzahl heterogener Akteure und Institutionen. „Intelligente“ Evaluation muss

<sup>57</sup> Schön/Rein (1994).

<sup>58</sup> Ähnlich wie die für die Umwelt-, aber auch die Technikpolitik entwickelten „Diskurs-“ und „Mediationsverfahren“; z. B. van den Daele (1997).

<sup>59</sup> Vgl. Kuhlmann et al. (1999).

<sup>60</sup> Vgl. Cuhls (1998); Cuhls/Kuwahara (1994).

<sup>61</sup> Vgl. z. B. Rip et al. (1995); Sundermann et al. (1999).

deshalb über simple input-output-Rechnungen (*value for money*) hinausgehen. Die Verwendung von Evaluationsverfahren als Lernmedium, das die divergierenden Perspektiven beteiligter Akteure in Wissenschaft, Industrie und Politik nicht leugnet, sondern unterschiedliche Interessen bewusst als konkurrierende Erfolgskriterien thematisiert, kann eine „reflexive Wende“ gegenüber älteren, häufig naiven rationalistischen Konzepten politischer Ge-

staltungsinitiativen einleiten. Umwälzende Richtungsänderungen in der Forschungs- und Innovationspolitik lassen sich auf diese Weise zwar nicht herbeiführen — dies kann nur durch (macht-)politische Entscheidung geschehen —, doch die praktische Umsetzung solcher Richtungsänderungen wird durch Moderation auf der Grundlage von Evaluationsergebnissen erleichtert: Sie unterstützt die Lernfähigkeit der Akteure.

## Literaturverzeichnis

- Airaghi, Angelo, Niels E. Busch, Luke Georghiou, Stefan Kuhlmann, Marc J. Ledoux, Anthony F. J. van Raan, José Viana Baptista* (1999): Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes. Report to the European Commission, DG XII, Evaluation Unit. Brussels/Luxembourg (Office for Official Publications of the European Communities).
- Airaghi, Angelo, Carsten Becker, Bernard Bobe, Luke Georghiou, Volker Kemper, Thoms Rönström, Norbert Wiegart* (1995): EUREKA Evaluation Report. Brussels.
- Alewel, Karl* (1988): Beurteilung der Leistungen von Hochschulen. In: *Daniel, Hans-Dieter, Rudolf Fisch* (Hrsg.): Evaluation von Forschung. Methoden, Ergebnisse, Stellungnahmen. Konstanz, 41–58.
- Becher, Gerhard, Stefan Kuhlmann* (1995) (eds.): Evaluation of Technology Policy Programmes in Germany. Boston, Dordrecht, London (Kluwer Academic Publishers).
- Block, Hans-J., Wilhelm Krull* (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, 427–437.
- Bobe, Bernard, Susanne Bühner, Raniero Chelli, Hanne Eriksen, Zehra Peksen, Maria Nedeva, François Sand* (1999): The Continuous and Systematic Evaluation of EUREKA. Annual Impact Report 1999. Brussels.
- Bozeman, Barry, Julia Melkers* (eds.) (1993): Evaluating R&D Impacts: Methods and Practice. Boston, Dordrecht, London (Kluwer Academic Publishers).
- Braun, Dietmar, Sybille Hinze, Bärbel Hüsing, Stefan Kuhlmann, Klaus Menrad, Viola Peter* (1997): Interdisziplinäre Zentren für Klinische Forschung in der Frühphase der Bundesförderung — Vergleichende Analyse. Stuttgart (Fraunhofer IRB Verlag).
- Brockhoff, Klaus* (1999): Forschung und Entwicklung. Planung und Kontrolle. München, Wien (R. Oldenbourg), 5. erg. u. erw. Auflage.
- Bührer, Susanne, Viola Peter, Dietmar Braun, Stefan Kuhlmann* (1999): Interdisziplinäre Zentren für Klinische Forschung — Kommunikation & Kooperation: Prozessevaluation der Fördermaßnahme „Interdisziplinäre Zentren für Klinische Forschung an den Hochschulkliniken“. Stuttgart (Fraunhofer IRB Verlag).
- Bürgel, Hans D., Christine Haller, Markus Binder* (1996): F&E-Management. München (Vahlen).
- Bundesministerium für Bildung und Forschung, BMBF* (1998): Bundesbericht Forschung. Faktenbericht 1998, Bonn.
- Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, BLK* (1999): Bericht der internationalen Kommission zur Systemevaluation der DFG und der MPG (Juni 1999) (56. Broschüre „Forschungsförderung in Deutschland“). Download unter URL <http://www.blk-bonn.de/papers/forschungsfoerderung.pdf>.
- Callon, Michel* (1992): The Dynamics of Techno-Economic Networks. In: *Coombs, R., P. Saviotti, V. Walsh* (eds.): Technological Change and Company Strategies: Economic and sociological perspectives. London et al. (Academic Press Limited), 72–102.
- Callon, Michel, Philippe Larédo, Philippe Mustar* (1995): La gestion stratégique de la recherche et de la technologie. L'évaluation des programmes. Paris (Economica).
- Campbell, David F.J., B. Felderer* (1997): Evaluating Academic Research in Germany. Patterns and Policies. Vienna (Institute for Advanced Studies) (Political Science Series, No. 48).
- Cozzens, Susan, Julia Melkers* (1997): Use and Usefulness of Performance Measurement in State Science and Technology Programs. In: *Policy Studies Journal*, 25, 2.
- Cuhls, Kerstin* (1997): Retrospektive auf die Technikvorausschau und die technologiepolitische Anwendungen des Delphi-Verfahrens in Japan. Heidelberg (Physica/ Springer, Reihe „Technik, Wirtschaft und Politik“, Bd. 29).

- Cuhls, Kerstin, T. Kuwahara (1994): Outlook for Japanese and German Future Technology. Comparing Technology Forecast Surveys. Heidelberg (Physica/Springer Publishers, Series „Technology, Innovation and Policy“, Vol. 1).
- Cunningham, Paul N., Luke Georghiou, Kate Barker, Stefan Kuhlmann, Guido Reger, C. Marciano da Silva, L. Henriques (1994): Analysis of Experience in the Use of Verifiable Objectives. Brussels/Luxembourg (Commission of the European Communities, MONITOR/SPEAR — EUR 15634 EN).
- Daniel, Hans-Dieter (1993): Guardians of Science. Fairness and Reliability of Peer Review. Weinheim u. a. (VCH).
- Dekker, Wisse, et al. (1991): Report of the EUREKA Assessment Panel. Brussels.
- Derlien, Hans-Ulrich (1976): Die Erfolgskontrolle staatlicher Planung. Eine empirische Untersuchung über Organisation, Methode und Politik der Programmevaluation. Baden-Baden (Nomos).
- Dreher, Carsten (1997): Technologiepolitik und Technikdiffusion. Auswahl und Einsatz von Förderinstrumenten am Beispiel der Fertigungstechnik. Baden-Baden (Nomos).
- Edquist, Charles (Hg.) (1997): Systems of Innovation. Technologies, Institutions and Organizations, London/Washington (Pinter)
- Ergas, Henry (1987): Does Technology Policy Matter? In: Guile, B.R., H. Brooks (eds.): Technology and Global Industry. Companies and Nations in the World Economy. Washington D.C. (National Academy Press), 191–245.
- Etzkowitz, Henry, Loet Leydesdorff (2000): The dynamics of innovation: from National Systems and „Mode2“ to a Triple Helix of university-industry-government relations. In: Research Policy 29, 109–123.
- European Commission (1999): Second European Report on S&T Indicators. Key Figures. Luxembourg (Office for Official Publications).
- Felt, Ulrike, Helga Nowotny, K. Taschwer (1995): Wissenschaftsforschung. Eine Einführung. Frankfurt/New York (Campus).
- Fraunhofer-Gesellschaft (1998): Systemevaluierung der Fraunhofer-Gesellschaft. Bericht der Evaluierungskommission. Download unter URL [ftp://192.76.176.135/eval\\_fhg.pdf](ftp://192.76.176.135/eval_fhg.pdf).
- Freeman, Chris (1987): Technology Policy and Economic Performance: Lessons from Japan, London (Pinter).
- Georghiou, Luke (1995): Research Evaluation in European National Science and Technology Systems. In: Research Evaluation, Vol. 5, No. 1, 3–10.
- Georghiou, Luke (1998): Issues in the Evaluation of Innovation and Technology Policy. In: Evaluation, Vol. 4 (1), 37–51.
- Gerpott, Torsten J. (1999): Strategisches Technologie- und Innovationsmanagement. Eine konzentrierte Einführung. Stuttgart (Schäffer-Poeschel).
- Gerybadze, Alexander, Meyer-Krahmer, F., Reger, G. (1997): Globales Management von Forschung und Innovation. Stuttgart (Schäffer-Poeschel).
- Gibbons, Micheal, C. Limoges, H. Nowotny, S. Schwartzman, P. Scott, M. Trow (1994): The new production of knowledge. The dynamics of science and research in contemporary societies. London et al. (Sage).
- Grimmer, Klaus, Jürgen Häusler, Stefan Kuhlmann, Georg Simonis (Hg.) (1992): Politische Techniksteuerung — Forschungsstand und Forschungsperspektiven. Opladen (Leske + Budrich).
- Grupp, Hariolf (1998): Foundations of the Economics of Innovation: Theory, Measurement and Practice. Cheltenham (Edward Elgar Publishing).
- Grupp, Hariolf, Kuntze, U., Schmoch, U. (1995): New Technology Indicators for the Evaluation of Research and Development Programmes. In: Becher, G. / Kuhlmann, S. (eds.): Evaluation of Technology Policy Programmes in Germany, Boton/Dordrecht/London: Kluwer Academic Publishers, 243–28
- Grupp, Hariolf, Ulrich Schmoch (1992): Wissenschaftsbindung der Technik. Panorama der internationalen Entwicklung und sektorales Tableau für Deutschland. Heidelberg (Physica, Wirtschaftswissenschaftliche Beiträge 69).
- Guba, Egon G., Yvonne S. Lincoln (1989): Fourth Generation Evaluation. Newbury Park et al.: Sage.
- Hellstern, Gerd-Michael, Helmut Wollmann (1984): Handbuch zur Evaluierungsforschung, Band 1. Opladen (Westdeutscher Verlag).
- Hornbostel, Stefan (1997): Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen (Westdeutscher Verlag).
- Jungmittag, Andre, Frieder Meyer-Krahmer, Guido Reger (1999): Globalisation of R&D and Technology Markets — Trends, Motives, Consequences, in: Meyer-Krahmer, F. (ed.), Globalisation of R&D and Technology Markets: Consequences for National Innovation Policies: Berlin (Physica-Verl.), 37–78 (Technology, Innovation, and Policy. Series of the Fraunhofer Institute for Systems and Innovation Research (ISI) Vol.9).
- Klodt, Henning (1995): Grundlagen der Forschungs- und Technologiepolitik. München (Vahlen).
- Kodama, Fumio (1995): Emerging Patterns of Innovation. Sources of Japan’s Technological Edge. Boston (Harvard Business School Press).
- Krull, Wilhelm, Frieder Meyer-Krahmer (1996) (eds.): Science and technology in Germany. London (Cartermill).
- Kuhlmann, Stefan (2000a): Evaluation in der Forschungs- und Innovationspolitik. In: Stockmann, Reinhard (Hg.):

- Evaluationsforschung, Leverkusen (Leske+Budrich) (im Erscheinen).
- Kuhlmann, Stefan* (2000b): Moderation von Forschungs- und Technologiepolitik? Evaluationsverfahren als „reflexives“ Medium. In: *Martinsen, Renate, Georg Simonis* (Hg.): Demokratie und Technik — (k)eine Wahlverwandtschaft? Leverkusen (Leske+Budrich), 303–332.
- Kuhlmann, Stefan* (1999): Politisches System und Innovationssystem in „postnationalen“ Arenen. In: *Grimmer, Klaus, Stefan Kuhlmann, Frieder Meyer-Krahmer* (Hg.): Innovationspolitik in globalisierten Arenen. Leverkusen (Leske+Budrich), 9–37.
- Kuhlmann, Stefan* (1998a): Politikmoderation. Evaluationsverfahren in der Forschungs- und Technologiepolitik. Baden-Baden (Nomos).
- Kuhlmann, Stefan* (1998b): Moderation of Policy-making? Science and Technology Policy Evaluation beyond Impact Measurement: the Case of Germany. In: *Evaluation*, Vol. 4, No. 2, 130–148.
- Kuhlmann, Stefan, Susanne Bühner* (1999): Evaluation von Forschungs- und Technologiepolitik. In: *Sundermann, Karsten, Stephan Bröchler, Georg Simonis* (Hg.): Handbuch Technikfolgenabschätzung. Berlin (Sigma), 237–249.
- Kuhlmann, Stefan, Doris Holland* (1995a): Evaluation von Technologiepolitik in Deutschland — Konzepte, Anwendung, Perspektiven. Heidelberg (Physica-Verlag).
- Kuhlmann, Stefan, Doris Holland* (1995b): Erfolgsfaktoren der wirtschaftsnahen Forschung. Heidelberg (Physica-Verlag).
- Kuhlmann, Stefan, Guido Reger* (1996): Technology-intensive SMEs: Policies Supporting the Management of Growing Technological Complexity. In: *Cannell, W., B. Dankbaar* (eds.): Technology Management and Public Policy in the European Union. Luxembourg/ Oxford (Office for Official Publications of the European Communities/ Oxford University Press), 73–102.
- Kuhlmann, Stefan, Patries Boekholt, Luke Georghiou, Ken Guy, Jean-Alain Héraud, Philippe Laredo, Tarmo Lemola, Denis Loveridge, Terttu Luukkonen, Wolfgang Polt, Arie Rip, Luis Sanz-Menendez, Ruud Smits* (1999): Improving Distributed Intelligence in Complex Innovation Systems. Brussels/Luxembourg (Office for Official Publications of the European Communities) (URL [http://www.isi.fhg.de/abtlg/ti/pb\\_html/final.pdf](http://www.isi.fhg.de/abtlg/ti/pb_html/final.pdf)).
- Kuntze, Uwe, Kurt Hornschild* (1995): Evaluation of the Promotion of R&D Activities in Small and Medium-sized Enterprises. In: *Becher, G. / Kuhlmann, S.* (eds.): Evaluation of Technology Policy Programmes in Germany, Boston/Dordrecht/London (Kluwer Academic Publishers), 33–54.
- Lundvall, Bengt-Åke* (Hg.) (1992): National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning, London (Pinter).
- Lundvall, Bengt-Åke, Susana Borrás* (1998): The globalising learning economy: Implications for innovation policy. Luxembourg (Office for Official Publications of the European Communities) (Targeted Socio-Economic Research).
- Martinsen, Renate, Georg Simonis* (Hg.) 1995: Paradigmenwechsel in der Technologiepolitik? Opladen (Leske+Budrich).
- Mayntz, Renate* (1980): Implementation politischer Programme. Empirische Forschungsberichte. Königstein (Verlagsgruppe Athenäum/Hain/Scriptor; Neue Wissenschaftliche Bibliothek).
- Mayntz, Renate* (1983): Implementation politischer Programme II. Ansätze zur Theoriebildung. Opladen (Westdeutscher Verlag).
- Meyer-Krahmer, Frieder* (1989): Der Einfluss staatlicher Technologiepolitik auf industrielle Innovationen. Baden-Baden (Nomos).
- Meyer-Krahmer, Frieder* (1997): Science-based Technologies and Interdisciplinarity: Challenges for Firms and Policy. In: *Edquist, Ch.* (ed.) (1997): Systems of Innovation. Technologies, Institutions and Organizations. London/Washington (Pinter), 298–317.
- Meyer-Krahmer, Frieder* (1999): Was bedeutet Globalisierung für Aufgaben und Handlungsspielräume nationaler Innovationspolitiken? In: *Grimmer, Klaus, Stefan Kuhlmann, Frieder Meyer-Krahmer* (Hg.): Innovationspolitik in globalisierten Arenen. Neue Aufgaben für Forschung und Lehre: Forschungs- und Technologiepolitik im Wandel. Leverkusen (Leske+Budrich), 35–65.
- Meyer-Krahmer, Frieder/ Uwe Kuntze*, (1992): Bestandsaufnahme der Forschungs- und Technologiepolitik. In: *Grimmer, K./ Häusler, J./ Kuhlmann, S./ Simonis, G.* (Hg.): Politische Techniksteuerung — Forschungsstand und Forschungsperspektiven, Opladen (Leske+Budrich), 95–118.
- Meyer-Krahmer, Frieder, Ulrich Schmoch* (1998): Science-based technologies: university-industry interactions in four fields. In: *Research Policy* 27, 835–851.
- Mowery, David C.* (1994): Science and Technology Policy in Interdependent Economies. Boston/Dordrecht/London (Kluwer Academic Publishers).
- Neidhardt, Friedhelm* (1988): Selbststeuerung in der Forschungsförderung. Das Gutachterwesen der DFG. Opladen (Westdeutscher Verlag).
- Nelson, Richard R.* (Hg.) (1993): National Innovation Systems: A Comparative Analysis, Oxford/New York (Oxford University Press).
- Niosi, Jorge* (ed.) (1999): The Internationalization of Industrial R&D, in: *Research Policy* (Special Issue), vol. 28, nos. 2–3, 107–336.

- OECD (1997) (ed.): Policy Evaluation in Innovation and Technology, Towards Best Practices. Paris.
- Ormal et al. (1993): Evaluation of EUREKA Industrial and Economic Effects. Brussels.
- Osborne, D., T. Gaebler (1993): Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector. New York (Plume).
- Ostry, S., Richard R. Nelson (1995): Techno-Nationalism and Techno-Globalism. Conflict and Cooperation. Washington (The Brookings Institution).
- Papaconstantinou, George, W. Polt (1997): Policy Evaluation in Innovation and Technology: an Overview. In: OECD (1997) (ed.): Policy Evaluation in Innovation and Technology, Towards Best Practices. Paris.
- Patton, Michael Q. (1997): Utilization-Focused Evaluation. The New Century Text. Thousand Oaks et al. (Sage).
- Rammert, Werner (1997): Auf dem Weg in eine post-schumpeterianische Innovationsweise. Institutionelle Differenzierung, reflexive Modernisierung und interaktive Vernetzung im Bereich der Technikentwicklung. In: Bieber, D. (Hg.): Technikentwicklung und industrielle Arbeit. Frankfurt/New York (Campus), 45–71.
- Rip, Arie (1990): Implementation and Evaluation of Science & Technology Priorities and Programs, in: Cozzens, Susan (ed.): The Research System in Transition. Boston/Dordrecht/ London: Kluwer Academic Publishers, 263–280.
- Rip, Arie, Th. J. Misa, J. Schot (eds.) (1995): Managing Technology in Society. The Approach of Constructive Technology Assessment, London/New York (Pinter).
- Röbbecke, Martina, Dagmar Simon (1999): Zwischen Reputation und Markt. Ziele, Verfahren und Instrumente von (Selbst)Evaluierungen außeruniversitärer, öffentlicher Forschungseinrichtungen. Berlin (Wissenschaftszentrum, WZB).
- Roobeek, Annemieke J. M. (1990): Beyond the Technology Race. An Analysis of Technology Policy in Seven Industrial Countries. Amsterdam et al. (Elsevier).
- Schmoch, Ulrich, Sybille Hinze, Gerhard Jäckel, N. Kirsch, Frieder Meyer-Krahmer, Gunnar Münt (1996a): The Role of the Scientific Community in the Generation of Technology. In: Reger, Guido, Ulrich Schmoch (eds.): Organisation of Science and Technology at the Watershed. The Academic and Industrial Perspective. Heidelberg (Physica/Springer, Series „Technology, Innovation, and Policy“, Vol. 3), 1–138.
- Schmoch, Ulrich, Sybille Breiner, Kerstin Cuhls, Sybille Hinze, Gunnar Münt (1996b): The Organisation of Interdisciplinarity — Research Structures in the Areas of Medical Lasers and Neural Networks. In: Reger, Guido, Ulrich Schmoch (eds.): Organisation of Science and Technology at the Watershed. The Academic and Industrial Perspective. Heidelberg (Physica/Springer, Series „Technology, Innovation, and Policy“, Vol. 3), 267–372.
- Schön, Donald, M. Rein (1994): Frame Reflection. Toward the Resolution of Intractable Policy Controversies. New York (BasicBooks).
- Scriven, Michael (1991): Evaluation Thesaurus, Newbury Park et al. (Sage).
- Shapira, Philip, Gordon Kingsley, Jan Youtie (1997): Manufacturing Partnerships: Evaluation in the Context of Government Reform. In: Evaluation and Program Planning, 2, 1, 103–112.
- Stockmann, Reinhard (Hg.) (2000): Evaluationsforschung, Leverkusen (Leske+Budrich) (im Erscheinen)
- Sundermann, Karsten, Stephan Bröchler, Georg Simonis (Hg.) (1999): Handbuch Technikfolgenabschätzung. Berlin (Sigma).
- van den Daele, Wolfgang (1997): Risikodiskussionen am „Runden Tisch“. Partizipative Technikfolgenabschätzung zu gentechnisch erzeugten herbizidresistenten Pflanzen. In: Martinsen, Renate (Hg.): Politik und Biotechnologie. Die Zumutung der Zukunft. Baden-Baden (Nomos), 281–301.
- van der Meulen, Barend, Arie Rip (1994): Research Institutes in Transition. Delft (Eburon).
- van Raan, Anthony F.J. (ed.) (1988): Handbook of Quantitative Studies of Science and Technology. Amsterdam (Elsevier).
- Widmer, Thomas (1996): Meta-Evaluation. Kriterien zur Bewertung von Evaluationen. Bern et al. (Haupt).
- Wollmann, Helmut (1998): Evaluation research and politics: Between a science-driven and a pluralist controversy-responsive policy-making model. Potential and limitations. Paper given to the Conference of the European Evaluation Society (EES). Roma, October.
- Worthen, Blaine R., J.R. Sanders, J.L. Fitzpatrick (1997): Program Evaluation. Alternative Approaches and Practical Guidelines (2nd ed.), White Plains, N.Y. (Longman).
- Zukunft nachgefragt. BMBF, Nr. 1, Juni 1998.

## **Performance Control and Learning Medium: Evaluation of Research and Innovation Policies**

### **Summary**

*The paper outlines the development of evaluation procedures in the field of science and innovation policy. After describing the general conditions of public research and technology support in Germany (section 2) we present an overview about the main concepts, application fields and methods of evaluation procedures within this policy area (section 3). The following sections sketch the ongoing structural change of research and innovation (section 4) and the consequences for the application of evaluation procedures in the political decision making process (section 5). The descriptions are illustrated by case examples of recent evaluation projects. The paper ends with some general conclusions (section 6).*