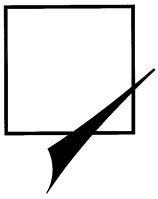
Moderation of Policy-Making?: Science and Technology Policy Evaluation Beyond Impact Measurement-The Case of Germany

Stefan Kuhlmann

Evaluation 1998; 4; 130
DOI: 10.1177/13563899822208491

Evaluation

[1356–3890 (199804) 4:2; 130–148; 004609]
Vol 4(2): 130–148



Moderation of Policy-Making?[1]

Science and Technology Policy Evaluation Beyond Impact Measurement—The Case of Germany

STEFAN KUHLMANN
Fraunhofer Institute Systems and Innovation Research (ISI), Karlsruhe

In the field of science and technology policies, for the most part, evaluation procedures are utilized as a way of measuring the scientific and technological quality or the socio-economic impacts of publicly funded research. Beyond this practice, could evaluation procedures be used as a medium for the 'moderation' of struggles, controversies and negotiations in the science and technology policy arena? The present article addresses this question using the German evaluation practice—moulded by a relatively high degree of institutional differentiation and autonomy of the major policy actors—as a background. After some theoretical considerations, a case study is presented illustrating the 'moderation approach': a multi-annual monitoring evaluation of eight newly created, publicly funded interdisciplinary clinical research centres at German university hospitals.

Introduction

Most industrialized countries now have at least some experience with the evaluation of science and technology (S/T) policies. For the most part, evaluation procedures are utilized as a way of measuring the scientific and technological quality or the (potential) socio-economic impacts of publicly funded research (Becher and Kuhlmann, 1995; Bozeman and Melkers, 1993; Georghiou, 1995; Rip, 1990). Beyond this practice, could evaluation procedures be used as a medium for the 'moderation' of struggles, controversies and negotiations in the S/T policy arena? The present article addresses this question using the German evaluation practice as a background.

The interest in the evaluation of S/T policies has increased markedly in Germany recently, mainly because of cut-backs in public budgets, which put considerable pressure on the prioritizing of financially effective state intervention. During the past two decades, a growing number of evaluation concepts were developed, the variety of methodologies was increased, and more and varied procedures were applied (Kuhlmann, 1995). The question remains whether this indicates a growing 'rationality' of the actors and decision-makers of the research system, or rather an increased sophistication of the players' battles for funds.

The German S/T system is characterized by a relatively high degree of institutional

differentiation and autonomy of the major actors (Krull and Meyer-Krahmer, 1996)—one could call it a 'multi-actor network'. Using evaluation procedures in such an environment is one means of creating more transparency in S/T policy. This implies an analysis of the frequently contradictory rationales of S/T policy players in order to facilitate a mutual critique and learning process. It also involves the assessment of indirect and unintended impacts of S/T policy initiatives in societal, economic and ecological spheres, and this would require a feedback of the knowledge gained through evaluation of the actor's networks and arenas.

In Germany, the authorities commissioning (external) evaluations of S/T programmes generally tend to be administrators with departmental orientations, and most of them do not, as the system stands, have a strong interest in studies with such far-reaching perspectives. There are, nevertheless, some signals of an intensified use of evaluation as part of a development strategy for S/T policies. This holds true in particular for those policy areas where a high modernization pressure on the one hand coincides with institutional and behavioural resistance on the other.

After some theoretical considerations, this article presents a case study illustrating this: a multi-annual monitoring evaluation of eight newly created, publicly funded interdisciplinary clinical research centres (ICRCs) at German university hospitals.

Policy Moderation and Evaluation: Theoretical Considerations

The centrality of government policy in societal development and for the problem-solving capacity of political-administrative systems is increasingly being called into question in empirical policy analyses as well as in theoretical debates. Ambitious attempts at 'policy planning' for societal modernization already failed in the 1970s. Empirical 'implementation research' (Mazmanian and Sabatier, 1981; Mayntz, 1983) demonstrated that this failure of state policy 'steering' is not merely caused by a lack of appropriate information to decision-makers, but by the contradictory nature and complexity of institutionally anchored 'frames' of action of the involved actors (Schön and Rein, 1994).

Today, policy-making is not only confronted by traditional conflicts of interest, but increasingly challenged by seemingly 'intractable' policy controversies (Hart and Kleiboer, 1995/96: 5). A typical example from many countries can be seen in the desperate efforts and repeating failures to modernize the national health care systems (e.g. Hicks, 1995/96: 3; see also Döhler, 1991). Depending on one's theoretical background, these settings could be characterized as an effect of incompatible societal communication codes (according to sociological systems theory, see Willke, 1992/96: 54) or as diverging institutional frames of actors 'in the sense that they represent mutually incompatible ways of seeing the policy situation' (Schön and Rein, 1994: 29; for neo-institutional approaches, see also DiMaggio and Powell, 1991).

If this action context is understood as a self-regulating 'negotiating system', questions must be faced about negotiating conditions or rules and—from the viewpoint of various actors—the relevant knowledge. Moderation becomes possible in the negotiating system if these rules and different actor perspectives are at least known to a moderator (an agent of the policy-administrative system). If the latter redistributes the

diverging bodies of knowledge within the negotiating system, it can provoke reflection and eventually facilitate 're-framing' (Schön and Rein, 1994).

I will restrict my conceptual considerations and empirical references to the research system in Germany, taken to mean a self-regulatory intermediary subsystem of the science system, the industry system and the political system. It can be shown how these highly differing spheres of society try to make their influence felt in the negotiating arena of the research system through their actors. Corporatist actors (research organizations, industrial associations and corporations, political-administrative institutions) negotiate the distribution of power (institutionalization and regulation) and the allocation of financial resources (research funds). Historically, the German research system represented a rather conventional corporatist consensus model (Mayntz and Scharpf, 1995: 24). Presently, it is increasingly challenged by deep structural changes and conflicts of interest, i.e. controversies tend to become intractable.

In this context, I will call *policy moderation* the targeted preparation and influencing (not unidirectional steering!) of science and technology policy-shaping and decision-making processes by policy administration actors. I am mainly interested in whether *evaluation procedures* can be designed as a *communication medium* for moderation processes and put to practical use in negotiating systems. The *main hypothesis* is that policy evaluation can, if it takes the heterogeneous perspectives (frames) of the participating actors into consideration, make a real contribution to the moderation of policies in negotiating systems. Government research policy administrators act as moderators, who can 'afford' independent evaluations of past, current or planned policy interventions in order to stimulate 'debates to facilitate decision-making' in negotiating systems. Such debates must not be designed exclusively to ease logrolling and to reach a consensus in any case, and decisions can also be made with parties dissenting, if the balance of power allows (for similar policy evaluation concepts, see Finne et al., 1995; de Laat, 1996). Ideally, moderated debates stimulate the emergence and the adoption of non-routinized, new perspectives on the subjects of negotiation.

Changing Research Systems and Science and Technology Policies

The 'innovation system' (e.g. Edquist, 1997; Keck, 1993; Lundvall, 1988), namely the functional cluster of industrial innovation activities, research institutions, education institutions and related political-administrative structures, is confronted with a great variety of challenges. They may be economic (e.g. globalization of markets and technological competition), socio-economic (e.g. employment systems, vocational training) or ecological (e.g. conservation of resources, sustainable development) in nature.

Research infrastructures and their networks must adapt to changed modalities of knowledge production in research and technology. Gibbons and colleagues summarize this development as the transition from 'Mode 1' to 'Mode 2':

> 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 characterized by homogeneity, Mode 2 by heterogeneity. Organizationally, Mode 1 is hierarchical and tends to preserve its form, while Mode 2 is more heterarchical and transient. (Gibbons et al., 1994: 3; see also Grupp, 1992, 1994; Schmoch et al., 1996)

This change corresponds with increased competition and, at the same time, increased networking and self-organization of research institutions. Rip (1994) calls this 'the emergence of a postmodern research system', 'becoming heterogeneous, in more than one sense'.

In Germany, the semi-autonomous and simultaneously interwoven actors in the S/T policy arena which jointly form the research system can be separated into three groups (see Figure 1).

The Public Scientific Research Facilities They have in common a 'science' based communication model and value system that governs their professional activities. They are further connected by the fact that they are completely or mainly financed by state funds (federal government, federal state or Länder, European Commission, or others). Well known differences in the form of public funding of the different types of facilities reflect varying research policy orientations of the institutions. Thus, the Max Planck Society and the German Research Association (DFG) are 100 percent financed by the central government and the federal states (Länder), and the universities, whose research (except from third-party sources) is mainly financed by the federal states, are characterized by self-organization within 'scientific communities'. The National Research Centres and government research institutes (departmental research) are very much more closely linked with the interests of the federal ministries financing them, but as a result of

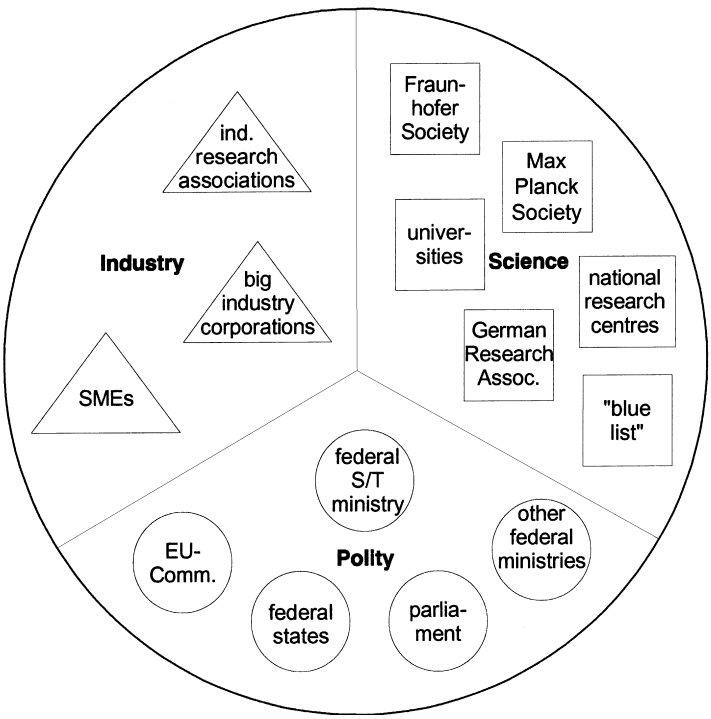


Figure 1. Actors in the S/T Policy Arena

their own institutional momentum and sheer size, can develop considerable own 'scientific-academic' interests, which diverge from those of their state financiers. The Fraunhofer Society, which receives only a relatively small basic funding from the state, as a contract research organization is to a large extent dependent on the interests of its customers in the research market. However, as a relatively large corporate actor, it also undeniably possesses clout and a certain say in the political arena.

Industrial Research Facilities Applied research and experimental development in industry have the largest share of research in Germany (especially in the electrotechnical, chemical and pharmaceutical industries). Even the development efforts of small and medium-sized enterprises—for example the machine tool industry—are relatively high in Germany, compared with other countries (Kuhlmann and Reger, 1996). In the S/T policy arena, large enterprises often appear as independent actors (Jansen, 1991); otherwise the Association of German Industry (BDI) acts for them while small and medium-sized enterprises are less organized.

The Policy Administration System This has become considerably differentiated within the past decades. From the 1960s, and especially in the 1970s, the federal Ministry of Research increasingly appeared on the scene together with the federal states (which are responsible for running the universities), offering generous technology support policies that stimulated the growth of non-university research organizations and formed ties with them (Stucke, 1993). Since the 1980s, some federal states have also been involved as initiators of own technology and innovation measures. Finally, since about 1985, the European Commission has also been playing an increasingly important part in the S/T policy field in Germany (Grande, 1995; Reger and Kuhlmann, 1995). It is often overlooked that, besides the Ministry for Research, other departments of the national government are actively involved in S/T policy not only as operators of research institutions, but also as initiators of independent research support programmes (e.g. the Federal Ministry of the Environment). The German parliament (Bundestag) also takes part in S/T policy decisions, although it has less opportunity to influence decisions than the legislatures of other countries, such as the USA.

In the S/T policy arena, there is no obvious centre of political power, and none of the actors dominates. All the actors in the arena are simultaneously—and most of them indeed primarily—members of their 'home system'. They belong to the science system, the industrial system or the political system, and the self-perception of the actors in the research system is clearly formed by the orientations and interests of their systems of origin. The S/T policy arena as such is not a permanent institution (with exclusive, self-referential codes), rather, it exists as an intermediary hybrid structure in precarious balance.

The Federal Ministry of Research can be characterized as the only genuine actor of this intermediary hybrid structure; it can be defined as the interface between science, industrial research and general state science and technology policy. This Ministry is inconceivable without the S/T policy arena, whereas the majority of the other actors in the S/T policy arena—if there were no longer a research system in the sense understood here—could revert back to their 'home bases' (i.e. pure science with public

funding; pure private industrial development, etc.). The Ministry of Research does not have a dominant role in the German research system—other actors (e.g. other government ministries and federal states) do not permit this. Nonetheless, the Ministry has a relatively strong status. This particular arrangement has evolved over time and only rarely allows the Ministry to target or strategically 'control' policy processes in the policy network of the research system. Nevertheless, the Ministry is able to exert significant influence both as a moderator within the research policy network and as an advocate of its own interests.

In the past decades, in keeping with its influential position, the Ministry of Research has determined the type and the selection of instruments utilized in S/T policy. The spectrum of instruments is large in principle (see Table 1): it reaches from the base funding of research facilities, via various forms of financial stimuli to carrying out research and experimental development in public or industrial research laboratories (technology and innovation programmes), up to the design of an 'innovation-orientated' infrastructure, including the institutions and mechanisms of technology transfer. These tools characterized the practice of S/T policy in the Federal Republic of Germany over the past three decades (Krull and Meyer-Krahmer, 1996), although funds have not been flowing so plentifully since the beginning of the 1990s[2] because, presently, state budgets are constrained.

Simple political attempts at top-down steering have little chance in this politically differentiated environment. S/T policy does not take place in a system of hierarchic decision-making, but in a multidimensional, networked negotiating system (Grande, 1995: 333; see also Mowery, 1994: 257). Furthermore, because of institutional differentiation (at the national, federal and European Commission level) the inner complexity of the policy administration system has grown in the recent past and the institutional conditions for governing have changed accordingly. Interest conflicts can easily become deadlocked.

Against this background, a transformation of the justification patterns and the assessment criteria for S/T policy is emerging. Actors in policy administration feel obliged to widen the legitimizing mechanisms beyond the familiar corporatist mediation of interests lobbying between scientific experts, industry and politics. 'Technology policy dialogue' is the new catchword, recalling to mind the 'Concerted Actions' of the late 1960s—a phase followed by significant socio-political modernization efforts in Germany. This illustrates the general observation

> That a re-framing is more likely to occur when the controversy is situated in an institutional context, where the price of inertia and protracted policy stalemate can be high . . . In addition, interdependent actors in policy networks cannot afford to jeopardize their interrelationships because of a conflict on issue X, when they find themselves less far apart or even potentially allied, against yet others on issues Y and Z that are to be resolved in the same network setting. (cf. Hart and Kleiboer, 1995/96: 9–10; see also Schön and Rein, 1994)

What does all this mean in terms of the function of evaluation? A successful moderation of generally cooperative strategies of actors who are nevertheless pursuing rival interests, requires 'intelligent' knowledge about the potential outcomes of different game strategies, and this possibly from the perspective of various actors: a challenge to conventional evaluation procedures!

Evaluation Results as Input to Policy Moderation

One can define evaluation as methodology-based analysis and assessment of the appropriateness of S/T policy assumptions and targets, of the related measures and their impacts, and of the goal attainment (Kuhlmann and Holland, 1995: 199; Kuhlmann and Meyer-Krahmer, 1995: 3). In order to design policy evaluations as a medium for moderation as well, some additional conceptual assumptions must be applied, which could be formulated as follows:

(1) The analyses and assessments in the framework of an evaluation procedure should take multiple actor perspectives into consideration, as regards both methodology and content, in order to be utilized in moderation processes. The different interests of participating actor groups and correspondingly differing perceptions of problems should explicitly be taken up conceptually and methodically.

Table 1. Public Research and Technology Policy Instruments

In a narrow sense	*In a broader sense*
(1) Institutional funding	(4) Public demand and purchasing
National Research Centres	(5) Corporatist measures
Max Planck Society	Long-term visions; technology foresight
Fraunhofer Society	Technology assessment
Higher Education Institutions	Technology Council
Others	Awareness initiatives
(2) Financial incentives	(6) (Continuing) education; training
Indirect promotion programmes (e.g. CIM)	(7) Public policy
Technology promotion programmes (cooperative R & D projects)	Competition policy
Risk capital	(De-)regulation
(3) Other innovation infrastructure and technology transfer mechanisms	Public stimulation of private demand
Information and consultancy for SMEs	
Demonstration centres	
Technology centres	
Cooperation, networks, people	

Meyer-Krahmer and Kuntze, 1992: 103.

(2) Evaluation procedures should be brought explicitly and visibly into the communication of negotiating systems and policy networks and, if possible, institutionalized as an iterative process element, functioning as a moderation medium depicting the actors' problem perceptions and reflecting their learning processes.

This leads to further conceptual and methodological requirements. Could evaluation procedures offer more than a kind of policy impact measurement? Probably yes—if one only looks far enough afield. Apart from the evaluation practices of research and technology policy—above all in the fields of social, education, communal and environmental policy programmes—evaluation research and evaluation procedures have developed considerably in the past 10 years (for a general overview, see Chelimsky and Shadish, 1997). Proceeding from the disappointing experience that evaluation results often produce only small impacts in political decision-making processes or, at the most, support one of the positions already represented in a policy arena, evaluation experts (and increasingly also policy-makers) tried to relax the fixed boundaries between evaluation and decision-making processes; indeed, partially even to integrate both spheres.

The key word of the new and broadened understanding of evaluation is 'negotiation' in actor arenas (Guba and Lincoln, 1989: 8). The result of the evaluations which were thus conceived is, by comparison with the conventional methodology, no longer '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 and Lincoln, 1989: 13). Here, therefore, the evaluation process, more exactly the communication in its course, steps into the foreground. The mediating character of the evaluation procedure becomes emphasized. One might even say that the medium becomes the message!

The new understanding of evaluation found rapid dissemination and recognition, particularly in the American 'evaluation community'. In 1993, the American Evaluation Association (AEA) devoted its annual conference to the subject 'Empowerment Evaluation'. The then AEA President, David Fetterman, defined empowerment evaluation as 'the use of evaluation concepts, techniques, and findings to foster improvement and self-determination' of evaluation participants (Fetterman et al., 1996: 4); evaluation should be understood as a medium for self-organized learning. The approach does not differ principally from Guba and Lincoln's concept, but Fetterman et al., however, worked out their theoretical approach far less systematically. The concept is clearly moulded by its use to increase the effectiveness of programmes to support disadvantaged social groups. In the meantime, a whole series of related evaluation concepts have been brought into circulation—besides 'empowerment evaluation' and 'fourth generation evaluation', 'participatory evaluation' and 'stakeholder-based evaluation' or 'collaborative evaluation'—so that some authors feel obliged to attempt classifications (e.g. Riquier, 1997: 5), which, however, reflect definitional rather than practical differences between evaluation practices.

For the concept of evaluation as a means to moderate S/T policy negotiations, the mechanisms of 'participatory evaluation' (for a summary, see Patton, 1997: 100) can

only be partly applied. Particularly misleading is the notion 'participant' when used in the context of the social policy programmes in the sense of 'beneficiaries', mostly members of socially disadvantaged groups, which are not characterized by strong, corporatistically organized actors.

To a certain extent, as shown by Bussmann (1996: 313) in the case of the corporatist political consensus culture in Switzerland, evaluation practice in policy-administrative negotiation systems can learn from the 'participatory evaluation' approach. The following aspects, especially, can be transferred or developed further out of the 'participative' approach:

(1) Evaluation could be conceived as a process of structured presentation and confrontation of (partly conflicting) actor perspectives.
(2) The evaluator could act as a facilitator, supporting the moderation of confrontations in the negotiating system by actors from the policy-administrative system.
(3) The evaluation aim is not only the assessment of facts from one individual actor's perspective (e.g. of the policy-administrative system), or the 'objective' testing of the appropriateness of a policy, but the stimulation of learning processes that can overcome deep-seated attitudes and positions.

These evaluation concepts aim above all to facilitate a 're-framing' (Schön and Rein, 1994) of the orientation of corporatist and policy-administrative actors. In this sense—this is an assumption—evaluation procedures could provide a decisive contribution to increase the 'rationality' of decision-making processes in corporatist and administrative negotiating systems and policy networks (see Figure 2). They can potentially achieve a systematization of context knowledge (using social-scientific instruments) for decision-making processes, especially a clarification of the actual or foreseeable impacts of completed, running or planned policies as they are perceived from the perspectives of various actors. In a multiple perspective framework, the expectation of an exact measurement of 'objective impacts' of a policy in the sense of immutable truth is neither possible nor wise. Instead, on the basis of analyses that acknowledge 'multiple objectives' concerning the relationship of policy aims, measures and impacts, a spectrum of alternative policy initiatives can be developed. These are policy alternatives with respect to different aims and differing approaches, target groups and time and factual ranges. Such multiple-perspective policy recommendations could be repeatedly fed into a negotiating system, where they help the actors to operate their strategic games more 'effectively', easing a re-framing of their perceptions of self-interest.

A key question here is whether and how the evaluators can cope with the conflicting demands, on the one hand, to remain sufficiently open to the various perspectives of the actors in the negotiating system and, on the other hand, to simultaneously maintain sufficient detachment towards individual interests, as well as the complete process, as expected of a 'neutral evaluator'. This question can only be answered case by case. We also face a structural problem here, that characterizes all 'mediating bodies', the control of which forms the basis for their institutional and professional success (e.g. consultants, lawyers and psychotherapists).

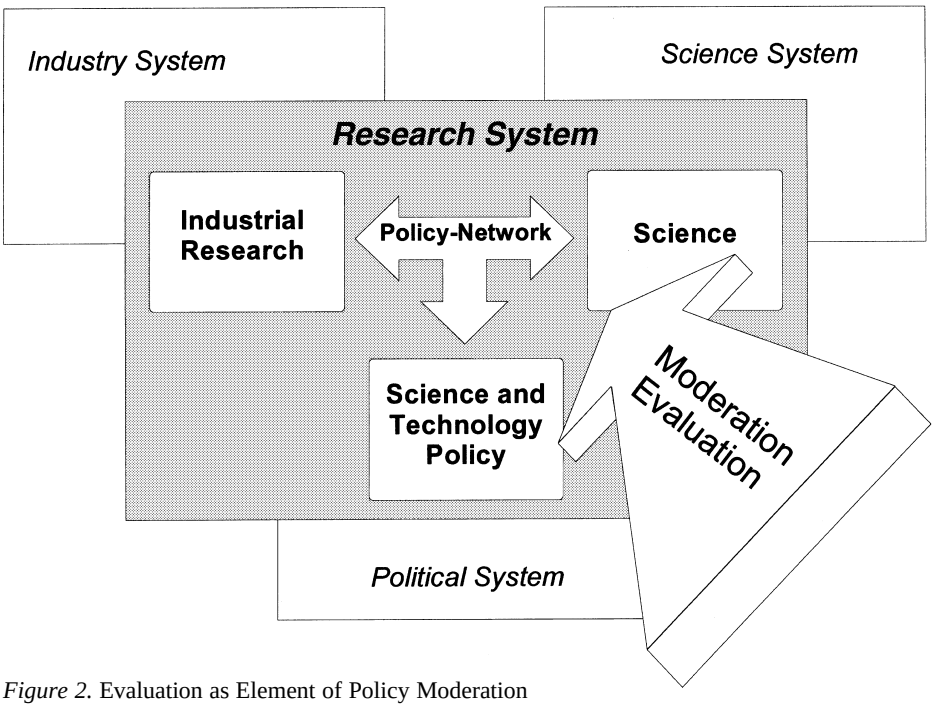


Figure 2. Evaluation as Element of Policy Moderation

Case study: Evaluation of Publicly Funded Clinical Research Centres

The Policy Initiative

The conditions for clinical research at German university hospitals are regarded as unfavourable and underdeveloped, by international standards (Braun, 1991: 46; 1992). In promoting interdisciplinary clinical research centres (ICRCs), the Federal Ministry for Education, Science, Research and Technology (BMBF) is seeking a lasting impetus to improve the situation. Within a competitive framework, eight universities were selected to establish pilot ICRCs in 1995/96. Federal funding is guaranteed for a certain time span (probably eight years), as 'tapered' start-up finance; the ICRCs are supposed to be funded in the medium-term mainly by their universities and the responsible federal state government. The main targets of the BMBF programme are:

(1) The establishment of efficient interdisciplinary clinical research structures.
(2) The development of specific research profiles of the participating university hospitals.
(3) Qualified scientific training conditions for young clinical researchers.
(4) Qualitative and competitive allocation of public research funds.
(5) Transparent financial management of research on the one hand and medical care on the other.

The concept of the ICRCs grants the chosen faculties and university hospitals a large degree of freedom in organization and decision-making, but combines this with demanding requirements for scientific performance, and also the development of innovative and effective management structures for clinical research. It is hoped that in the course of the programme various more or less effective ICRC structures will emerge which can serve as models for other university hospitals. A fundamental function of the BMBF promotional measure is to disseminate experiences with the new structures beyond the eight participating universities to the actor network involved in clinical research, in order to provide an impulse for the further development of the whole system in Germany. It is expected that, eventually, clinical research, which has been neglected in favour of patient care and student training, should regain its position as the key discipline between fundamental research and practice.

Policy Arena of Clinical Research

All stakeholders in the arena of clinical research (including the evaluators) are challenged by the fact that the majority of German medical university faculties are conservative and hierarchically structured institutions that do not have much experience—to a great extent are not even interested—in modern interdisciplinary (i.e. clinical) research. The interests represented in this arena often obstruct each other.

From the perspective of a newly created ICRC, the university hospital and its departments and units constitute the most important actor context that is mainly committed to the health system. At the same time, the clinic overlaps personnel-wise and organizationally to a great extent with the medical faculty of the university, which has to follow the orientations of the education and science systems.

The task of patient care is in the allocation of resources (money, personnel and physical resources) and in the organization of a clinic (relatively independent hospitals and departments, staffing schedules, accounting procedures, authority to issue instructions, etc.). The second major task follows from the teaching obligations of the medical faculty. The share of the apportioned workload between patient care and teaching by members of a clinical centre is an important ‘bargaining point’ in the context of the ICRCs’ development. The relative ‘strength’ and the interlinking of the research activities of the promoted ICRCs with institutions of fundamental research in other areas of the university or outside must also be carefully watched.

In recent years, and to an increasing extent, the actors of the policy-administrative system have come to play an active role in the arena of health research policy. They do not, however, pursue uniform or even coherent aims. From the perspective of a clinical research centre, the BMBF of course plays a central role as financier. The Ministry of Science of the local federal state also has decisive influence, which is why the commitment and the attempts at reform differ greatly from state to state. The Federal Ministry for Health, which as a regulator tries to advance the reform of the health system, also exercises indirect influence in the background.

In this arena, there is no one dominant actor who could simply ‘force through’ the establishment of effective ICRCs in the university hospitals. In terms of the creation of favourable conditions for clinical research, the confrontation of interests in the policy arena can be characterized as nearly ‘intractable’. Only when several of the central

actors make a cooperative effort to set up and run a centre does this have a certain chance of success. The decision of several actors to establish an ICRC and the application for supporting funds in the framework of the BMBF programme can be regarded as a cooperative initiative in this sense.

Aims of the Evaluation

This diagnosis was the intellectual starting point of a comprehensive evaluation effort. In 1995/6, a multi-annual monitoring evaluation was started, designed as a continuous learning and 'moderation' process between ICRCs, the funding authorities (BMBF and regional science ministries), and the independent evaluators.

The German medical faculties enjoy a high degree of autonomy. It is up to them whether they actively support the new ICRCs or let them develop in a rather hostile environment. The evaluation process is supposed to look from an outsider's perspective into the arena and to provide 'objective' information on contested actor interests and on the factors hampering and fostering the ICRC's development. The evaluators are expected to debate this information with the ICRCs, to feed the results obtained into the 'negotiating arena' on a regular basis, thus finally helping the ICRCs to survive and to develop.

In particular, the evaluation project has (1) to analyse the actual development of the ICRCs in relation to the programme's targets; (2) to compare the achievements of the centres against the background of their specific local (clinical, scientific, infrastructural, financial, regulative) conditions; and (3) to put forward recommendations for the future development of ICRCs. In parallel, the evaluation team is supposed (4) to ensure an open dialogue with the centres in order to ease the empirical analysis and to feed back the analytical results; (5) to actively support a working group of the ICRCs' speakers and other leading representatives through information inputs, in order to raise the quality and focus of debates; and (6) to present and discuss intermediate and final evaluation results regularly with all participating actors.

Evaluation Concept

As guiding principle, we assume that the success of a (research) facility depends in essence on the ability and the commitment of the individuals involved, but that given structural frame conditions decisively influence their freedom of action. The organizational structure, the division of tasks among the clinical centres, time management, and especially the allocation of resources (financial, personnel and technical), belong under this category. A main question is therefore: *Are the selected regulations, procedures and organizational structures of an ICRC capable of asserting the academic and clinical modernization targets of the overall programme?*

Clinical centres are not autonomous organizations but are closely linked to their environment. They do not have a large pool of personnel at their disposal solely for the centre; on the contrary, the majority of the positions are filled with persons who are also employed in the medical faculties and the university clinics. Conflicts of interest can arise here. A second main question arises: *Which strategies do clinical centres apply in order to reconcile the conflicting interests of their members (such as the allocation of*

responsible positions, authority to make decision-making competencies and resources) with the aims of the centre?

Methodologically speaking, the monitoring evaluation concentrates on before/after comparisons of the individual centres and on comparisons of the developments of the centres with each other. The following target dimensions for centre development were compiled on the basis of the promotion aims and discussions with experts and taken as a basis for the empirical investigations:

The Centre Should be Administratively Anchored to the University, the Faculty and the Clinic This evaluation dimension covers the whole centre concept. By this we mean the hierarchical and procedural organization of the clinical centre, its links with clinic and faculty, its cooperation and communication with external actors, and its management culture. Possible indicators are:

- interest constellations
- justification and suitability of the selected organizational model
- representativeness of committees (including procedures, elections, etc.)
- administration (personnel, material)
- possibilities for communication
- actual linking of labs and wards (proximity)
- access of researchers to patients.

Efficient Research Structures This evaluation dimension covers the long term strategical orientation of the research, the focus of the research activities on particular themes as well as mechanisms for a lasting reinforcement of the clinical research. Possible indicators are:

- thematic profiling of research
- motivation of the centre researchers
- independent decision-making structures
- mechanisms (structures) for coordinating research efforts
- settlement of time resources/working time models
- distribution of resources (material/personnel).

Interdisciplinarity and Raising the Level of the Scientific Quality This evaluation dimension contains the interlocking of theoretical and clinical subjects as well as of basic and applied research. Possible indicators are:

- cooperation
- interdisciplinarity (as precondition for project promotion)
- stimuli, motivation
- sanctions
- internal selection and quality control.

Support for Young Clinical Researchers This evaluation dimension comprises the mechanisms established in and around the centre aimed at training junior clinical researchers and to offer junior scientists help in their careers. Possible indicators are:

- existence, autonomy and success of junior staff research groups
- part-time release from patient care duties to participate in clinical research
- postgraduate and post-doc programmes and scientific achievements of junior staff (publications, prizes).

Transparent Financing This evaluation dimension covers mechanisms to make research, teaching and patient care in the university clinic and medical faculty, the sources of funding of the clinical centre, and the organization of the long term redistribution of resources in favour of clinical research more transparent. Possible indicators are:

- separate financing and book-keeping for research and patient care
- internal project control mechanisms.

The evaluation procedure covers several research steps in all: (1) comparative analysis of the participating clinical centres in the early phase (1996) and in the fourth year of the project; (2) monitoring analysis of the clinical centres in the course of the BMBF programme; and (3) utilization of the evaluation as a 'medium of communication' with the clinical research centres.

The participating centres were analysed as case studies. In the summer and fall of 1996, intensive interviews were conducted in the individual centres. To this end, an interview guideline was elaborated, which was then adapted to each specific centre location and to the specific position or function of each interviewee. Between 10 and 15 in-depth interviews took place in each centre. Interviewees were representatives of the board of directors of the centre, leaders of the research groups, members of the scientific teams, including junior researchers and representatives of the near environment of the ICRCs (Dean of the Faculty, Administrative Director of the hospital and scientists not employed by the centre). The case studies served as a basis for a thorough comparative analysis (see Braun et al., 1997). As of winter 1996/97, intermediate results on the state of the development of the ICRCs were available.

Intermediate Results

The BMBF intervention has certainly brought a 'fresh breeze' into the research landscape of the German university hospitals and medical faculties. It stimulated central actors everywhere in the domain to grapple with the aims of the promotional measure, i.e. structural reforms—either as an applicant (successful or not) who had to incorporate the programme's aims in the planning for a new centre, or even as a formally uninvolved faculty member, who will, however, have to live in the future with the existence of an ICRC.

A relatively extensive process of application and appraisal of the clinical research centres by external scientific peer panels has pushed performance-based fund allocation procedures into the world of clinical research. All actions and decisions of the applicants were affected by the final recommendation of these external panels, which thus directly and indirectly acted like a 'yeast' fermenting performance-orientation and modernization in clinical research. Many actors willingly support process certainly. It is therefore all the more important for them to see the problem constellations, which could impede further developments, as carefully, objectively and free of emotions as possible.

The following six paragraphs give a flavour of intermediate results emerging from the evaluation.

(1) In selecting possible organizational forms to anchor clinical research in Germany, the favourite was a research association model, based on voluntary collaboration, stimulated by financial, scientific and career incentives. This model appears to correspond best to the legal requirements and to the existing constellations of actors in the German medical faculties.

(2) A greater problem is to contribute to the creation of a common ICRC identity of the researchers on the operative level. Their constant switching between the medical faculty and the centre helps to diffuse the centre idea, but hardly encourages identification with its structural aims. For this it would be necessary to try to increase the researchers' motivation during their active ICRC participation. The beginnings of structures designed to promote communication and cooperation can already be found, especially in those places where in the past cross-project and inter-institute research cooperations were already part of everyday routine.

(3) In the present early phase of the programme, the centres have complied with the formal requirement of interdisciplinary cooperation by interdisciplinary project design and by the composition of the project teams. Further steps to encourage inter-disciplinary discussions in current work, beyond the project boundaries, and cooperation with the environment outside the centre have hardly begun and could be developed much further.

(4) The centres offer their members incentives to constructive collaboration because of their financing, their interdisciplinary design and the (expected) first-class quality of their research. Additional, explicit incentive mechanisms (e.g. bonus/reward systems) have not been developed, nor have formal sanction mechanisms in case of violation of the centre's objectives.

(5) Research time (sabbatical leave) for young staff members and the financing of rotating jobs has been handled very differently in the participating centres to date. It remains to be seen how young staff members who are interested in research can be best integrated. Some centres have initiated ambitious measures for the interdisciplinary scientific education of junior researchers, whereas others have still to take decisive initiatives in this direction.

(6) The structures of research financing in the clinics and medical faculty are presently going through a time of upheaval; some have already made very good progress towards establishing transparent models of research financing, while others are presently doing so, and yet others are adopting a wait-and-see position. The provision of the universities' own share of the financing of the clinical research centres was formally pledged by the university hospitals and medical faculties, but the fulfilment of this pledge will not proceed smoothly everywhere.

The discussion process between the ICRCs, the promoters and the evaluators, intensified in January 1997 with the presentation of these intermediary evaluation results. The two very different worlds of clinical research and of social science-based policy evaluation

tried to start joint discussions and even collided for a time (see also Smith et al., 1997). But, eventually the results were 'fed in' to the negotiation system and were thereby understood as a contribution to the discussion and not as a final judgement on the clinical research centres. The centres were very much still in a *status nascendi*; in this stage of the complex process of institutionalization of modern, interdisciplinary research organizations, quick assessments on the basis of simple, easily 'calculable' indicators would have been inadequate. The comparison informed the centres about their common institutional traits and differences and provided an overview of the different development paths which had been taken by the centres. Only at a later date will it be possible to compare these paths and their positive and negative effects.

The evaluative analysis and comparison of the centres made it easier for all participants—the centres, associated organizations and the programme sponsors—to distinguish between specific or local problems (such as particular infrastructures or the character of the local negotiating system) and those difficulties caused by general issues facing clinical research in Germany. In this way a 're-framing' of the controversial actor perspectives was facilitated and opportunities for learning processes were opened up. The contemporary concern in Germany is about the survival and development of institutionalized first-rate clinical research. Although institutional conditions are difficult and in part adverse, it seems unlikely that the momentum for the ongoing modernization of clinical research in Germany will come to a halt.

The monitoring evaluation has fulfilled its function as a medium for moderation in the first phase of the promotional process, in that it used the cooperative orientation mobilized by the promotional measure of at least some of the actors in the negotiating system, observed their activities and problems in establishing the centres, relayed these observations back to the negotiating system and thus created pre-requisites for a re-orientation of actor perspectives.

Outlook

The evaluation of the ICRC initiative—which was deliberately designed as a policy moderation experiment—represents one of only a few cases, that stands apart from the general German S/T (policy) 'evaluation culture' (Kuhlmann, 1995). Moreover, this innovative evaluation effort is still in its *status nascendi*, and it remains to be seen whether it will be ever successful.

However, there is little doubt that the underlying structure of the S/T policy arena and the present and near-future policy agenda call for 'advanced forms of mediated policy dialogue, and the conduct of policy evaluations producing actionable knowledge for policy-makers instead of mechanistic judgements about their own performance' (Hart and Kleiboer, 1995/96: 10). Related procedures and practices require further development and testing.

Ongoing analyses of the practical use of various evaluation concepts indicate that there is an important relationship between the nature of the related policy negotiation arena and the extent to which evaluation can be used as a moderation medium. If a policy is designed as a limited programme and its targeted actor network is only vaguely specified, then the arena will also remain undetermined, and an evaluation

process cannot fulfil an explicit moderation function. Even in these circumstances, it is not impossible that the evaluation, by dint of the information presented, will bring enlightenment to the policy-administrative system and even to the target audience, and by supporting learning, can contribute to the 'fine tuning' of initiatives.

If, on the other hand, a political initiative is aimed at the establishment or further development of institutional structures, and if the actors aimed at are clearly identifiable, then the conditions obtained are more favourable for the functioning of an evaluation process as a moderating medium, even if initially only a small measure of willingness to cooperate and reach consensus is present in the negotiating system.

Notes

1. This paper is an extended version of a contribution to the Topical Interest Group 'Evaluation of Research, Technology and Development' during the Annual Meeting of the American Evaluation Association (AEA), Atlanta, 6–7 November 1996; and a contribution to the 1997 Conference of the European Evaluation Society (EES), Stockholm, March 1997.
2. The total expenditure on research amounted to 36 billion dollars in Germany in 1994. Measured against the GDP, Germany is still in the top group of the leading industrialized countries in this respect (2.9 percent in 1989; 2.3 percent in 1994), although it fell back in the past years by comparison with Japan (3 percent in 1989; 2.9 percent in 1993) and the USA (2.8 percent in 1989; 2.5 percent in 1994) (BMBF, 1996).

References

Becher, G. and S. Kuhlmann (eds) (1995) *Evaluation of Technology Policy Programs in Germany*. Boston, MA: Kluwer Academic Publishers.

Bozeman, B. and J. Melkers (eds) (1993) *Evaluating R & D Impacts: Methods and Practice*. Boston, MA: Kluwer Acadamic Publishers.

Braun, D. (1991) *Die Einflußmöglichkeiten der Forschungsförderung auf Strukturprobleme der Gesundheitsforschung in der Bundesrepublik, Vol. 15*. Bonn: Schriftenreihe zum Programm der Bundesregierung Forschung und Entwicklung im Dienste der Gesundheit.

Braun, D. (1992) *Probleme und Perspektiven der Gesundheitsforschung in den Vereinigten Staaten, Frankreich und England, Vol. 23.* Bonn: Schriftenreihe zum Programm der Bundesregierung Forschung und Entwicklung im Dienste der Gesundheit.

Braun, D., S. Hinze, B. Hüsing, S. Kuhlmann, K. Menrad and V. Peter (1997) *Vergleichende Analyse Klinischer Forschungszentren in der Frühphase der Bundesförderung. Zwischenbericht zur Prozeßevaluation 'Interdisziplinäre Zentren für Klinische Forschung an den Hochschulkliniken'*. Stuttgart: Fraunhofer IRB Verlag.

Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (BMBF) (1996) *Bundesbericht Forschung*. Bonn: BMBF.

Bussmann, W. (1996) 'Democracy and Evaluation's Contribution to Negotiation, Empowerment and Information. Some Findings from Swiss Democratic Experience', *Evaluation* 2(3): 307–19.

Chelimsky, E. and W. R. Shadish (eds) (1997) *Evaluation for the 21st Century, A Handbook*. Thousand Oaks, CA: Sage.

DiMaggio, P. J. and W. W. Powell (1991) 'Introduction', in W. W. Powell and P. J. DiMaggio (eds) *The New Institutionalism in Organizational Analysis*, pp. 2–38. Chicago, IL: The University of Chicago Press.

Döhler, M. (1991) 'Policy Networks, Opportunity Structures and Neo-Conservative Reform Strategies in Health Policy', in B. Marin and R. Mayntz (eds) *Policy Networks. Empirical Evidence and Theoretical Considerations*, pp. 235–96. Frankfurt: Campus; Westview Press.

Edquist, Ch. (ed.) (1997) *Systems of Innovation. Technologies, Institutions and Organizations.* London: Pinter.

Fetterman, D. M., S. J. Kaftarian and A. Wandersman (eds) (1996) *Empowerment Evaluation. Knowledge and Tools for Self-Assessment & Accountability.* Thousand Oaks, CA: Sage.

Finne, H., M. Levin and T. Nilssen (1995) 'Trailing Research. A Model for Useful Program Evaluation', *Evaluation* 1(1): 11–31.

Georghiou, L. (1995) 'Research Evaluation in European National Science and Technology Systems', *Research Evaluation* 5(1): 3–10.

Gibbons, M., C. Limoges, H. Nowotny, S. Schwartzman, P. Scott and M. Trow (1994) *The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Grande, E. (1995) 'Regieren in verflochtenen Verhandlungssystemen', in R. Mayntz and F. W. Scharpf (eds) *Gesellschaftliche Selbstregelung und politische Steuerung*, pp. 327–68. Frankfurt: Campus, Westview Press.

Grupp, H. (ed.) (1992) *Dynamics of Science-Based Innovations*. Berlin: Springer.

Grupp, H. (1994) 'Technology at the Beginning of the 21st Century', *Technology Analysis & Strategic Management* 6(4): 379–409.

Guba, E. G. and Y. S. Lincoln (1989) *Fourth Generation Evaluation.* Newbury Park, CA: Sage.

Hart, P. and M. Kleiboer (1995/96) 'Policy Controversies in the Negotiatory State', *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization* 8(4): 5–25.

Hicks, E. (1995/96) 'Forethoughts', *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization* 8(4): 3–4.

Jansen, D. (1991) 'Policy Networks and Change: The Case of High-Tec Super Conductors', in R. Mayntz and B. Marin (eds) *Policy Networks. Empirical Evidence and Theoretical Considerations*, pp. 137–74. Frankfurt: Campus; Westview Press.

Keck, O. (1993) 'The National System for Technical Innovation in Germany', in R. R. Nelson (ed.) *National Innovation Systems. A Comparative Analysis*, pp. 115–57. Oxford: Oxford University Press.

Krull, W. and F. Meyer-Krahmer (1996) 'Science, Technology, and Innovation in Germany—Changes and Challenges in the 1990s', in W. Krull and F. Meyer-Krahmer (eds) *Science and Technology in Germany*, pp. 3–29. London: Cartermill.

Kuhlmann, S. (1995) 'Patterns of Science and Technology Policy Evaluation in Germany', *Research Evaluation* 5(1): 23–33.

Kuhlmann, S. and D. Holland (1995) *Evaluation von Technologiepolitik in Deutschland—Konzepte, Anwendung, Perspektiven, vol. 12.* Heidelberg: Physica-Verlag.

Kuhlmann, S. and F. Meyer-Krahmer (1995) 'Practice of Technology Policy Evaluation in Germany: Introduction and Overview', in G. Becher and S. Kuhlmann (eds) *Evaluation of Technology Policy Programs in Germany*, pp. 3–29. Boston, MA: Kluwer Academic Publishers.

Kuhlmann, S. and G. Reger (1996) 'Technology-Intensive SMEs: Policies Supporting the Management of Growing Technological Complexity', in W. Cannell and B. Dankbaar (eds) *Technology Management and Public Policy in the European Union*, pp. 73–102. Oxford: Oxford University Press.

de Laat, B. (1996) 'Scripts for the Future. Technology Foresight, Strategic Evaluation and Socio-Technical Networks: The Confrontation of Script-Based Scenarios', PhD thesis, University of Amsterdam.

Lundvall, B. A. (1988) 'Innovation as an Interactive Process: from User–Producer Interaction to the National System of Innovation', in G. Dosi, C. Freeman, R. Nelson and G. Silverberg (eds) *Technical Innovation and Economic Theory*. London.

Mayntz, R. (1983) *Implementation Politischer Programme II. Ansätze zur Theoriebildung*. Opladen: Westdeutscher Verlag.

Mayntz, R. and F. W. Scharpf (1995) 'Steuerung und Selbstorganization in staatsnahen Sektoren', in H. G. Dieselben (ed.) *Gesellschaftliche Selbstregelung und Politische Steuerung*, pp. 9–38. Frankfurt: Campus, Westview Press.

Mazmanian, D. A. and P. Sabatier (eds) (1981) *Effective Policy Implementation*. Lexington, MA: Lexington Books.

Meyer-Krahmer, F. and U. Kuntze (1992) 'Bestandsaufnahme der Forschungs- und Technologiepolitik', in K. Grimmer, J, Häusler, S. Kuhlmann and G. Simonis (eds) *Politische Technik-steuerung—Forschungsstand und Forschungsperspektiven*, pp. 95–118. Opladen: Leske and Budrich.

Mowery, D. C. (1994) *Science and Technology Policy in Interdependent Economies*. Boston, MA: Kluwer Academic Publishers.

Patton, M. Q. (1997) *Utilization-Focused Evaluation. The New Century Text.* Thousand Oaks, CA: Sage.

Reger, G. and S. Kuhlmann (1995) *European Technology Policy in Germany*. Heidelberg: Physica-Verlag.

Rip, A. (1990) 'Implementation and Evaluation of Science & Technology Priorities and Programs', in S. Cozzens, P. Healey, A. Rip and J. Ziman (eds) *The Research System in Transition*, pp. 263–80. Boston, MA: Kluwer Academic Publishers.

Rip, A. (1994) *The Postmodern Research System*, paper presented at the European Research Conference in 'Il Ciocco' (November), Tuscany; and at the European Science Forum on the History of Science Policy in Europe (December), Florence.

Riquier, G. (1997) *Stakeholders' Participation in Public Policy Evaluations: Impact on Organizational Learning*, paper presented at the Annual Conference of the European Evaluation Society (EES) (March), Stockholm.

Schmoch, U., S. Breiner, K. Cuhls, S. Hinze and G. Münt (1996) 'The Organization of Interdisciplinarity—Research Structures in the Areas of Medical Lasers and Neural Networks', in G. Reger and U. Schmoch (eds) *Organization of Science and Technology at the Watershed. The Academic and Industrial Perspective*, pp. 267–372. Heidelberg: Physica-Verlag.

Schön, D. and M. Rein (1994) *Frame Reflection: Toward the Resolution of Intractable Policy Controversies*. New York: Basic Books.

Smith, A., D. Preston, D. Buchanan and S. Jordan (1997) 'When Two Worlds Collide. Conducting a Management Evaluation in a Medical Environment', *Evaluation* 3(1): 49–68.

Stucke, A. (1993) *Institutionalisierung der Forschungspolitik. Entstehung, Entwicklung und Steuerungsprobleme des Bundesforschungsministeriums*. Frankfurt: Campus, Westview Press.

Willke, H. (1992/96) *Ironie des Staates. Die Grundlinien einer Staatstheorie polyzentrischer Gesellschaft*. Frankfurt: Suhrkamp.

STEFAN KUHLMANN is a political scientist and head of the department 'Technology Analysis and Innovation Strategies' at the Fraunhofer Institute Systems and Innovation Research (ISI), Breslauer Strasse 48, D-76139 Karlsruhe. He is (co-)author of several books and articles on science and technlogy policy evaluation. [email: sk@isi.fhg.de]