Chapter 25

EVALUATION AS A MEDIUM OF SCIENCE & TECHNOLOGY POLICY: RECENT DEVELOPMENTS IN GERMANY AND BEYOND

by

Dr. Stefan Kuhlmann
Fraunhofer Institute Systems and Innovation Research (ISI), Karlsruhe, Germany

Introduction

In Germany, the interest in the evaluation of S&T policies has increased markedly in recent years, mainly because of cut-backs in public budgets which put considerable pressure on the prioritising of financially effective state intervention. During the past two decades a growing number of evaluation concepts have been developed, the variety of methodologies has increased, and more and varied procedures have been applied (Kuhlmann, 1995). The question could be raised as to whether this indicates a new awareness of the need for efficacy and efficiency, or rather increased sophistication in the players’ scramble for funds; a practical example at the end of this chapter will show that – not surprisingly – the answer can be found half-way between these two extremes.

This chapter will: i) discuss the need for advanced S&T policy planning instruments; ii) take a brief look at current S&T policy evaluation practice in Germany; iii) develop some theoretical considerations about policy making in multi-actor governance structures; iv) sketch the outline of a related evaluation concept; v) illustrate this concept by a recent evaluation project; and vi) provide a European outlook.

The changing research system and S&T policies

Is there a need for “mediated”, more “reflexive” S&T policy making, including evaluation concepts that go beyond the measurement of certain indicators as a basis for political decisions?

At the beginning of the 21st century, the “innovation system” (Edquist, 1997; Keck, 1993; Lundvall, 1988) – i.e. the functional cluster of industrial innovation activities, research system, education system and related policy-administrative structures – is, in Germany as well as in most industrialised countries, confronted with a wide variety of challenges, be they economic (e.g. globalisation of markets and technological competition), socio-economic (e.g. employment systems, vocational training) or ecological (e.g. conservation of resources, sustainable development).
The research system must adapt to changes in knowledge production in research and technology. Gibbons et al. (1994) summarise this development as the transition from “Mode 1” to “Mode 2”. 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 et al., 1994, p. 3; Rip, 1992; Grupp, 1992 and 1994; Schmoch et al., 1996).

The semi-autonomous and simultaneously interwoven actors in the S&T policy arena which together form the German research system can be separated into three groups (see Figure 1):

◊ **Industrial research facilities.** Applied research and experimental development in industry have relatively the largest share of research in Germany (especially in the electrotechnical, chemical and pharmaceutical industry). The development efforts of even small and medium-sized enterprises - for example the machine tool industry - are relatively high in Germany compared with other countries (Kuhlmann and Reger, 1996). Large enterprises often appear in the research and technology policy arena as independent actors (Jansen, 1991); otherwise the Association of German Industry (BDI) acts for them. Small and medium-sized enterprises seldom act in a strongly organised fashion in the research and technology policy arena.

◊ **Public scientific research facilities.** These have in common the reference to “science” as their communication model and value system of their professional activities; they are further linked by the fact that they are completely or mainly financed by state funds (federal government, federal state or Länder, European Commission, others). Well-known differences in the form of public funding of the different types of facilities reflect varying research policy orientations of the institutions: the Max Planck Society and the German Research Society (DFG), which are 100 per cent 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 characterised by an orientation to self-organisation 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 can through institutional momentum and sheer size, 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 organisation, is to a large extent dependent on the interests of its customers in the research market. However, as a relatively large corporatistic actor it also undeniably possesses clout and a certain say in the political arena.

◊ **The actors of the policy administration system** have become considerably differentiated within the past decades. Seen from an historical perspective, in the 1960s the federal Ministry of Research appeared increasingly on the scene besides the Länder and, particularly in the 1970s by means of generous promotion policies, stimulated the growth of research organisations and formed ties with them (Stucke, 1993). Since the 1980s some of the federal states have also been involved as actors on the research and technology policy scene. In addition, since about 1985, the European Commission has also been playing an increasingly important role in the research and technology policy field in Germany (Reger and Kuhlmann, 1995; Grande, 1995). To a certain extent, in addition to the Ministry for Research, other
departments of the federal government and, of course, the German parliament (Bundestag), also take part in research and technology policy decisions.

None of these actors dominates the German S&T policy arena – there is no obvious centre of political power. All the actors 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 (in a narrow sense). The way the actors see themselves in the S&T policy arena is clearly a product of the orientations and interests of their systems of origin. The S&T policy arena as such has not yet become a deeply rooted, permanent institution (with exclusive, auto-referential codes), rather it exists as an intermediary hybrid governance structure in endangered balance.

The most genuine actor of this intermediary governance structure is by far the Federal Ministry of Research, acting as an interface between science, industrial research and general state science and technology policy. The existence of this Ministry would be inconceivable in the absence of the S&T policy arena, whereas the majority of the other actors - if the governance structure as it is understood here was to lose its influence - could revert back to their “home bases” (i.e. pure science with public funding; pure private industrial development). The special role of the Ministry of Research in the research system does not assign it a dominant position - the other actors (e.g. other government ministries, federal states) would not permit this - but does confer on it a relatively strong status. This peculiar constellation has emerged through historical evolution; it seldom allows the Ministry targeted, strategic “control” of the policy processes in the policy network of the research system, but it does enable it to exert a formative influence as a moderator – not, of course, a neutral one, rather a partisan of its own institutional interests.
In keeping with its influential position, until recently the Ministry of Research determined the type and the selection of instruments utilised in S&T policy. The spectrum of instruments is wide (see Figure 2): it reaches from the base funding of research facilities, via various forms of financial stimuli to conducting research and experimental development in public or industrial research laboratories (technology and innovation programmes), to the design of an “innovation-oriented” infrastructure, including the institutions and mechanisms of technology transfer. These tools characterise 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 quite so plentifully since the beginning of the 1990s – nowadays state budgets suffer from emaciation.

![Figure 2. Public research & technology policy instruments](image)

<table>
<thead>
<tr>
<th>In a narrow sense</th>
<th>In a broader sense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Institutional funding</td>
<td>4. Public demand and purchasing</td>
</tr>
<tr>
<td>• National Research Centres</td>
<td>5. Corporatist measures</td>
</tr>
<tr>
<td>• Max Planck Society</td>
<td>• Long-term visions; technology foresight</td>
</tr>
<tr>
<td>• Fraunhofer Society</td>
<td>• Technology assessment</td>
</tr>
<tr>
<td>• Higher Education Institutions</td>
<td>• “Technology Council”</td>
</tr>
<tr>
<td>• Others</td>
<td>• Awareness initiatives</td>
</tr>
<tr>
<td>2. Financial incentives</td>
<td>6. (Continuing) education; training</td>
</tr>
<tr>
<td>• Indirect promotion programmes (e.g. CIM)</td>
<td></td>
</tr>
<tr>
<td>• Technology promotion programmes (co-operative R&amp;D projects)</td>
<td></td>
</tr>
<tr>
<td>• Risk capital</td>
<td></td>
</tr>
<tr>
<td>3. Other innovation infrastructure and technology transfer mechanisms</td>
<td>7. Public policy</td>
</tr>
<tr>
<td>• Information and consultancy for SMEs</td>
<td>• Competition policy</td>
</tr>
<tr>
<td>• Demonstration centres</td>
<td>• (De-) Regulation</td>
</tr>
<tr>
<td>• Technology centres</td>
<td>• Public stimulation of private demand</td>
</tr>
<tr>
<td>• Co-operation, networks, people</td>
<td></td>
</tr>
</tbody>
</table>


In this politically scattered environment German S&T policies cannot be viewed as hierarchic decision making, but rather as a multi-dimensional, multi-actor negotiating system (Grande, 1995, p. 333; Mowery, 1994, p. 257). Furthermore, due to institutional differentiation (at the national, federal and European levels), the complexity of the policy administration system continues to grow and the institutional conditions for governing have accordingly changed significantly. Conflicts of interest can often become insoluble controversies and end in deadlock, so simple political attempts at top-down steering have absolutely no chance of success.

Against this background, the actors in policy administration feel obliged to widen the legitimising mechanisms beyond the familiar corporatistic mediation of interest lobbying among scientific experts, industry and politics (Mayntz and Scharpf, 1995) - “technology policy dialogue” is the new catchword. What does this mean for evaluation? Successfully moderating the co-operative strategies of actors who are pursuing rival interests requires “intelligent” knowledge of the potential outcomes of the different strategies, and this from the perspective of all the various actors – a real challenge to conventional evaluation procedures!
The present state of evaluation in Germany

The practical use and functions of evaluation procedures in the German S&T system are wide-ranging (cf. for the following Kuhlmann, 1995), and the expectations and requirements of the evaluation users vary significantly:

◊ from legitimising the distribution of public money and demonstrating effective use of funding;
◊ via targeting and “controlling” in the sense of improved management and “fine tuning” of S&T policy programmes;
◊ to an attempt to improve transparency in the rules of the game and the profusion of research funding and subsidies, and to enhance the information basis for shaping S&T policies, in the sense of acting as a moderator between the diverging and competing interests of the various players in the S&T system.

From an historical point of view, we can differentiate two basic lines of development in the field of S&T evaluation in Germany, as in other industrialised countries. They can be sketched as a shell model:

◊ The “core” consists of the peer review procedures, later completed by the measurement of research performance of single researchers or groups (using bibliometrics, etc.) as science-internal instruments for the allocation of funds to research institutions. These instruments are widely used in the German S&T system, in particular in the field of basic and long-term applied research.
◊ The “outer layer” of the shell consists of evaluation studies in the sense of impact analyses of S&T policy programmes. This concept can be seen as science-external; it tends to be initiated by policy makers to prove the achievement of politically set scientific, technological, economic or societal targets. Use of impact analyses has spread since the late 1970s, in particular in the areas of innovation and strategic technology programmes.

Evaluation of research and research institutions

The appropriateness of measuring and evaluating research performance has repeatedly been debated in Germany (Buchholz, 1995; Daniel and Fisch, 1990; Backes-Gellner, 1989). Fundamental questions include: Can the performance of science be measured and evaluated at all? Who can judge the importance and value of scientific knowledge? Peer reviews are considered to be an important “instrument of self-regulation of science” (Daniel, 1993). Peer review is the principle evaluation instrument of the research funding mechanism of the Deutsche Forschungsgemeinschaft (DFG). The DFG plays a central role in funding basic research at universities, distributing its budget (DM 1 820 million in 1994) in the form of grants to university researchers; these grants are grouped along several programme lines. A significant part of the budget is spent as individual project grants (Normalverfahren). Grant applications are evaluated by honorary peers who are elected on a four-year rota by the entire scientific community.

Over the level of single research projects, there is as yet no regular evaluation practice in Germany covering the research performance of university institutions or scientific disciplines, although a number of different evaluation efforts have been made. An important variation from this
general picture is provided by the Wissenschaftsrat, a science policy advisory body set up in 1957 to advise the federal and state governments on all matters of higher education and research policy (Block and Krull, 1990); today it has 54 members representing academia and government. The Wissenschaftsrat prepares reports and recommendations on the structural development of universities and research institutes.

Over the past few years several novel attempts have been made to evaluate the performance of industrially oriented state-supported R&D institutions: in 1993-94 the Federal Ministry for Research and Technology initiated two industry-led, panel-type evaluations, one of the national research centres and another of the most relevant information technology labs. In both cases representatives of important industrial companies were given the opportunity to assess the industrial orientation of the researchers’ work and the (potential) attractiveness of their research results for industrial application. Another novel concept was initiated by the government of a federal state: the Ministry for Economic Affairs of Baden-Württemberg supported the development of a generally applicable, qualitative and in-depth evaluation concept for the 35 industrially oriented R&D institutes that are co-funded by this Ministry. The approach is based on “success factors”, helping the managers of the R&D institutes to run their lab effectively, and on “performance criteria”, helping the funding government bodies to assess the actual fulfilment of the tasks of the institutes (Kuhlmann and Holland, 1995a).

In summary, German evaluation practice with regards the performance of research and research institutions can be characterised as strong and at the same time somewhat scattered: strong because of its high degree of self-organisation which guarantees consensus and commitment amongst the researchers; scattered, since to date institutionally oriented evaluation efforts have not been organised under a broader conceptual umbrella aiming at a strategic re-orientation of the S&T system.

S&T programme evaluation

Since the 1970s, more and more public funds for S&T have been spent on project funding, mostly in the framework of targeted policy programmes (usually of five years’ duration); the BMBF currently allocates more than 40 per cent of its budget to programmes. Accordingly, evaluation studies, in the sense of impact analyses of programmes, have been carried out increasingly since the end of the 1970s, usually performed by outside, independent research institutes on behalf of the policy administration system. This experience of nearly two decades of programme evaluation has helped to establish an “evaluation scene” in Germany, consisting of a group of experts and institutes in the fields of economics and social sciences and covering a relatively broad range of methodological approaches and evaluation instruments (Becher and Kuhlmann, 1995).

In 1992-93 the former German Ministry for Research and Technology (BMFT) initiated a “meta-evaluation” of the experience gained (Kuhlmann and Holland, 1995b; Kuhlmann, 1995). The task was to suggest improvements in the methodological instruments and the management of programme evaluations of the BMFT on the basis of a comparative analysis and assessment of the methods and results of 50 programme evaluations commissioned by the Ministry since 1985. In order to maximise compatibility in future evaluation activities, a “basic pack” of minimal requirements for the implementation of evaluation studies was recommended, concerning the planning of the evaluation procedure, the selection of evaluation teams, the content, scope, depth and extent of the analysis, the methods and indicators, the presentation and use of results.
Policy moderation and evaluation: theoretical considerations

One could question whether and how this basic evaluation experience in Germany could be further developed in order to better meet the challenges of changing innovation systems and political conditions for science, technology and innovation. The answer to this question should be based on a functional model of policy making in modern differentiated societies.

The central function of state policy in societal development and for the problem-solving capacity of policy-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 modernisation failed in the 1970s. Empirical “implementation research” (Mazmanian and Sabatier, 1981; Mayntz, 1983) demonstrated that the failure of state policy “steering” is not merely a problem of a lack of appropriate information by the actors, but stems from the contradictory nature and complexity of the institutionally anchored “frames” of their action (Schön and Rein, 1994).

Today, policy making is not only confronted with traditional conflicts of interest but is increasingly challenged by seemingly “intractable” policy controversies (Hart and Kleiboer 1995-96); a typical example can be seen in the desperate efforts and repeated failures of many countries to modernise their national health-care systems (Döhler, 1991). Depending on one’s theoretical background, such settings could be characterised as an effect of incompatible communication codes of differentiated societal sub-systems (according to sociological systems theory; Willke 1992-96) 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, p. 29; DiMaggio and Powell, 1991, for neo-institutional approaches).

If this context of action is understood not as a restriction or even an impasse, but rather as an auto-dynamic “negotiating system”, questions about negotiating conditions, rules and - from the viewpoint of various actors - the relevant knowledge crop up. Moderation becomes possible in the negotiating system if these rules and the perspectives of the different actors are known at least to a moderator (an agent of the policy-administrative system): if the moderator redistributes the diverging bodies of knowledge within the negotiating system, he can confront actors with the “frame-multiplicity”, provoke reflection and eventually facilitate a “re-framing” (Schön and Rein, 1994) of the actors’ perspectives.

Considering the “research system” in Germany - taken to mean an auto-dynamic sub-set of the science system, the industry system and the political system - we can see how these highly differing spheres of society attempt to make their influence felt in the negotiating systems of the research system through their actors and to make arrangements with science and technology policies. Historically, the German S&T policy arena and related governance structure represented a rather conventional corporatist consensus model (Mayntz and Scharpf, 1995), while today it is increasingly challenged by deep structural changes and interest conflicts, i.e. controversies tend to become intractable.

In this context policy moderation by policy administration actors is more likely to occur than unidirectional steering of science and technology policy. Units of the research policy administration system 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 reach a consensus in any case, and decisions can also be made with parties dissenting, if the balance of power allows.
Evaluation results as input to policy moderation

The above-mentioned meta-evaluation defined 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 goal attainment (Kuhlmann and Holland, 1995b, p. 199; Kuhlmann and Meyer-Krahmer, 1995). In order to design policy evaluations as a medium for moderation as well, some additional conceptual assumptions must be applied, which could hypothetically be formulated as follows (for the following see also Kuhlmann, 1997):

◊ 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 utilised in moderation processes;
◊ evaluation procedures should be brought explicitly and visibly into the negotiating systems and policy networks and if possible institutionalised as an iterative process element, functioning as a moderation medium depicting the actors’ perceptions of problems and reflecting their learning processes.

But how can evaluation procedures perform more than a kind of policy impact measurement? If one looks far enough afield - outside the evaluation practices of S&T policy - evaluation research and evaluation procedures have experienced considerable development in the past ten years, above all in the fields of social, education, communal and environmental policy programmes: Proceeding from the disappointing experience that evaluation results often produce only small impacts in political decision making or, at the most, support one of the positions already represented in the policy arena, evaluation experts (and increasingly also policy makers) are trying 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” (Guba and Lincoln, 1989, p. 8). The results of evaluations conceived in this light are, 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, p. 13). Here therefore the evaluation process, and more exactly the communication in the course of the process, is crucial. Once the mediating character of the evaluation procedure emphasized – the medium becomes the message!

The concepts of “empowerment evaluation”, “fourth generation evaluation”, “participatory evaluation”, “stakeholder-based evaluation” or “collaborative evaluation” were rapidly disseminated and recognised, particularly in the American “evaluation community” (Fetterman et al., 1996). They are clearly moulded by their task to increase the effectiveness of programmes to support disadvantaged social groups. For the purpose of evaluation as a medium to moderate S&T policy negotiations, the mechanisms of “participatory evaluation” (see e.g. the summary by Patton, 1997, p. 100) can only be partly applied. Particularly misleading is the notion of “participant” - used in the sense of “beneficiaries”, who are mostly members of socially disadvantaged groups which are not characterised by strong, corporatistically organised actors.

To a certain extent, however, as shown by Bussmann (1996, p. 313) with the example of the clearly corporatistic political consensus culture in Switzerland, evaluation practice in policy-administrative negotiation systems can learn from the “participatory evaluation” approach. In particular, the following aspects can be transferred or further developed from the “participative” approach:
evaluation could be conceived as a process of structured presentation and confrontation of (partly conflicting) actor perspectives;

◊ the evaluator could act as a facilitator, he supports the moderation of confrontations in the negotiating system between actors from the policy-administrative system;

◊ the evaluation aim is not only the assessment of facts from an individual actor’s perspective (e.g. of the policy-administrative system), or the “objective” testing of the appropriateness of a policy, but rather the stimulation of learning processes by overcoming ingrained actor orientations.

Figure 3. Evaluation as an element in policy moderation

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 evaluation procedures could provide a decisive contribution to increasing the “rationality” of the decision-making processes in corporatist and administrative negotiating systems and policy networks (Figure 3). They can achieve a systematisation of context knowledge (using social-scientific instruments) for decision-making processes, in particular a clarification of the actual or foreseeable impacts of completed, ongoing or planned policies as they are perceived from the perspectives of the various actors. In the framework of a multiple perspective concept, the expectation of an exact measurement of the “objective impacts” of a policy in the sense of immutable truth is no longer possible, or indeed wise. Instead, on the basis of so-to-speak “multiple objective” analyses concerning the relationship of policy aims, measures and impacts, a spectrum of alternative policy initiatives can be developed, as
alternatives with respect to different aims and differing approaches, target publics and time and factual ranges.

An example: evaluation of publicly funded clinical research centres

Several recent or ongoing S&T policy programme evaluation projects contain some elements of this type of evaluation concept, e.g. a recently started, multi-annual monitoring evaluation of newly created interdisciplinary clinical research centres at university hospitals, funded by the BMBF (for details see Kuhlmann, 1997; Braun et al., 1997).

The policy initiative: From an international comparison, the conditions for clinical research at German university hospitals are regarded as unfavourable and underdeveloped (Braun, 1991; Braun, 1992). By promoting interdisciplinary clinical research centres (ICRCs), the Federal Ministry for Education, Science, Research and Technology (BMBF) aims to provide a lasting impetus to improve this situation. Through a competition, eight universities were selected to establish pilot ICRCs in 1995-96. Federal funding is guaranteed for a period of time (probably eight years) as degressive kick-off financing; in the medium term, the ICRCs should be funded mainly by their universities and the federal state government responsible. The main targets of the BMBF programme are:

◊ the establishment of efficient interdisciplinary clinical research structures;
◊ the development of specific research profiles for the participating university hospitals,
◊ qualified scientific training conditions for young clinical researchers;
◊ qualitative and competitive allocation of public research funds;
◊ transparent financial management of research, on the one hand, and medical care, on the other.

The ICRC concept confers on the promoted faculties and university hospitals a large degree of freedom in organisation and decision making, but combines this with demanding requirements for scientific performance, but also for 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 network of actors involved in clinical research in order to provide impulses for the further development of the whole system. It is hoped that clinical research, which had tended to be neglected in favour of patient care and student training, will regain its ranking as a key discipline between fundamental research and medical practice.

Policy arena and governance structure of clinical research: The stakeholders in the arena of clinical research (and the evaluators) are challenged by the fact that the majority of German university medical faculties are conservative and hierarchically structured institutions that have little experience - and, 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 their departments and units form the most important actor context; mainly committed to the logic of the health system. At the same time, the clinic overlaps from a personnel and organisational point of view 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 most clearly distinguished; it is found in the allocation of resources (money, personnel, physical resources) and in the organisation of a clinic (relatively independent individual hospitals and departments, structure of the staffing schedule, accounting procedure, authority to issue instructions, etc.). In second place in the structure of interests come the teaching obligations of the medical faculty. The share of the apportioned workload in the area of patient care, as well as in the area of teaching, by members of a clinical centre is an important “bargaining point” in the context of ICRC development. The most sensitive aspects of the BMBF initiative are the relative “strength” and the interlinking of the interdisciplinary research activities of the promoted ICRCs with institutions of fundamental research in other areas of the university or outside.

The various actors of the policy-administrative system play an active role in the arena of health research policy, and in past years to an increasing extent. However, they do not pursue uniform, or even coherent, aims. From the perspective of a clinical research centre the BMBF plays a central role as financier. The Ministry of Science of the local federal state also has decisive influence, so that the commitment and the attempts at reform differ very much 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 single dominant actor who could simply “force through” the establishment of effective ICRCs in university hospitals. In terms of creating favourable conditions for interdisciplinary clinical research, the confrontation of interests in the policy arena can be characterised as nearly “intractable”. Only when several of the central actors make a co-operative effort to set up and run a centre, does the project 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 co-operative initiative in this sense.

Aims and concept of the evaluation: This diagnosis was the intellectual starting point of an encompassing evaluation effort. In 1995-96, a multi-annual monitoring evaluation was started, designed to act as a continuous learning and “moderation” process between ICRCs, the funding authorities (BMBF, regional science ministries) and the independent evaluators.

The German medical faculties enjoy a high degree of autonomy. It is up to them to decide whether they will actively support the new ICRCs or leave them to develop in a rather hostile environment. The monitoring evaluation process is designed to look into the arena from an outsider’s perspective and to provide “objective” information on the conflicting interests of the actors and on factors fostering or hampering the development of the ICRCs. The role of the evaluators is to debate this information with the ICRCs, to feed the results back into the “negotiating arena” of competing and contentious actors, thus finally helping the ICRCs to survive and develop.

In particular, the evaluation project should: i) analyse the actual development of the ICRCs in relation to the programme’s targets; ii) compare the achievements of the centres against the background of their specific local (clinical, scientific, infrastructural, financial, regulative) conditions; and iii) make recommendations for the future development of ICRCs. In parallel, the evaluation team is supposed to: iv) ensure an open dialogue with the centres in order to facilitate the empirical analysis and feed back the analytical results into the ICRCs; v) actively support a working group of ICRC speakers and leading representatives through information inputs, in order to raise the quality and problem orientation of the debates; and vi) regularly present and discuss intermediate and final evaluation results with all actors.
The evaluation procedure comprises several stages of research: i) comparative analysis of the participating clinical centres in the early phase (1996) and in the fourth year of the project; ii) monitoring of the clinical centres in the course of the BMBF programme (1995 onwards); and iii) utilisation of the evaluation as a “medium of communication” in the environment of the clinical research centres (1995 onwards). In the beginning, the participating centres were analysed in the form of case studies. In Summer and Fall 1996 intensive interviews were conducted in the individual centres. The case studies served as a basis for a thorough comparative analysis. As of Winter 1996-97, intermediate results on the state of development of the ICRCs were available.

Intermediate results: The analysis results “fed in” to the negotiation system so far were to be understood as a contribution to the discussion rather than a final judgement on the situation of the clinical research centres. In 1996-97 the centres were very much in a status nascendi; at this stage of the complex process of institutionalisation of modern, interdisciplinary research organisations, quick assessments on the basis of simple, easily “calculable” indicators would have been inappropriate. The comparison informed the centres of their common institutional traits and differences, and provided an overview of the diverging paths of development which they had taken. Only at a later date will it be possible to compare the strengths and weaknesses of these different development paths.

The evaluative analyses of the centres and their comparison make it easier for all participants - the centres, their network of actors and the promoters of the programme - to distinguish between specific local problems (infrastructure; characteristics of the local negotiating system) and difficulties arising from the general situation of clinical research in Germany. In this way a re-framing of controversial perspectives was facilitated and opportunities for learning processes opened up. The question now is what are the chances of survival and development of institutionalised first-rate clinical research. The general feeling is that current institutional conditions are difficult, and in some cases even adverse, but that the process of lasting modernisation of clinical research which has been initiated in Germany will not be stopped.

The monitoring evaluation has fulfilled its function as a medium for moderation in the first phase of the promotional process, in that it has used the co-operative orientation mobilised 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 so created pre-requisites for a re-orientation of the actor perspective.

The European outlook

Compared to the complexity of the partly contradictory needs of changing research systems, the political institutions of the German innovation system – as well as related structures in other European countries and the European Commission – still have difficulties in linking up systematically with the changing requirements of multi-actor and multi-level policy-negotiating systems (and vice versa).

Using evaluation procedures, in the sense demonstrated above, as one source of creating transparency in S&T policy “games in multi-actor networks” would imply an analysis of the frequently contradictory rationales of S&T policy players in order to facilitate mutual critique and learning processes. This would also involve the assessment of indirect and unintended individual and interactive impacts of S&T policy measures in societal, economic and ecological spheres. In
principle, impact evaluations of ongoing or completed S&T policy initiatives could even be combined with and enriched through:

◊ *S&T foresight* efforts (Cameron *et al.*, 1996), providing debatable “visions” of various more-or-less desirable futures; and

◊ *technology assessment* procedures (Rip *et al.*, 1995), attempting to anticipate the potential positive or negative impacts of technological developments, and to feed back related information into the technology creation process.

Given such preconditions, policy evaluation could be seen as a formative element for future policy strategy development in multi-actor arenas of S&T policy networks. This type of advanced evaluation effort would act as an infusion of intelligence into the bargaining and coping strategies of the S&T policy players: an “Advanced Science and Technology Policy Planning (ASTPP) (1996-98)” Network (supported by the European Union’s TSER programme) is currently investigating how and to what extent policy impact evaluations, technology foresight and technology assessment can be used and further enhanced in order to make the interfaces between the research system, industry, society and polity more “intelligent”. By “intelligent” the ASTPP network means that the interrelations between S&T, industrial efforts, societal needs and political interventions (at the national and European levels) become more transparent, thus facilitating interactive collaboration. “*Intelligent* policy planning procedures would provide information that:

◊ is based on analyses of the *dynamics of changing research systems*, changing innovation processes, changing functions of public policies;

◊ is aware of *diverging* “frames” of actors’ perceptions;

◊ helps to formulate diverging perceptions of (even contentious) subjects in a more “*objective*” way, offering appropriate indicators and information processing mechanisms; and

◊ can be deployed through “*discussions*” between contesting actors (or between those that represent their views).

Under these conditions, policy evaluation efforts would go far beyond S&T performance measurement and contribute to the development and stabilisation of a democratic culture in S&T negotiating systems.
NOTE

1. Total expenditure on research in Germany amounted to 36 billion dollars in 1994. Measured against the GDP, Germany is still in the top group of the leading industrialised countries in this respect (1989: 2.9 per cent; 1994: 2.3 per cent), although it fell back in the past years by comparison with Japan (1989: 3.0 per cent; 1993: 2.9 per cent) and the USA (1989: 2.8 per cent; 1994: 2.5 per cent (BMBF, 1996).
REFERENCES


BUNDESMINISTERIUM FÜR BILDUNG, WISSENSCHAFT, FORSCHUNG UND TECHNOLOGIE (BMBF) (1996), Bundesbericht Forschung, Bonn.


