6 Future Policy Instruments: Evaluation of the Socio-Economic Effects in the European Research Area

Luke Georghiou and Stefan Kuhlmann

6.1 INTRODUCTION

This chapter addresses the evaluation of new policy instruments. The development of evaluation practice has tended to mirror the evolution of technology and innovation policy, moving from an initial (and ongoing) focus on collaborative RTD programmes in the 1980s and gradually shifting towards measures intended to enhance the environment for innovation and technology transfer (Georghiou, 1998). Most recently there has been an increasing interest in policies designed to build research and innovation capacity, encompassing human capital and mobility enhancement, infrastructures and the building of networks. This has been accompanied by a shift in the rationale for innovation policy, or at least an extension from the market failure arguments developed in the 1960s and applied strongly in the 1980s. The structuralist-evolutionist approach now recognises that while information failures and lack of appropriability of returns may cause an under-investment in RTD, they do not necessarily guide the policymaker to the most appropriate actions. Such guidance may be obtained from the systems of innovation approach, which tends to highlight the absence of bridging institutions and the need to overcome firms resistance to adopt new technologies. A full review of developments in economic rationales and their implications for evaluation practice is available in the report of the ASIF Project (Georghiou, Rigby and Cameron (eds), 2002).

The focus of this chapter is upon the challenge to evaluation presented by the emergence of a series of innovative policy instruments under the European Research Area concept. At the same time consideration is given to policy measures emerging at national level.

6.2 GENERAL DESCRIPTION

Following a document published by the Commissioner for Research at the beginning of 2000, *Towards a European Research Area* (European Commission, 2000), the European Commission has been implementing a major change in its research policies. For the past fifteen years, its two principal instruments for research funding have been programmes of collaborative research and work in the Commission's own laboratories. The intention is to move away from these two instruments, which account for a small fraction (around 5%) of public funding, and to mobilise the entire research resource of Europe.

The principal aim of the ERA measures is to reinforce European competitiveness or to contribute to the solution of important societal problems through the mobilisation of a critical mass of research and technological development resources and skills existing in Europe. However, it also entails closer integration of research policies, which has implications both for policy implementation and subsequent evaluation.

In line with this aim, new policy instruments have been developed for application within and beyond the sixth Framework Programme. Key amongst these policy instruments are Networks of Excellence, Integrated Projects and provision for participation in Member States' programmes under Article 69.

6.2.1 Networks of Excellence

The objective of Networks of Excellence (NoE) is to reinforce European scientific and technological excellence by means of a progressive and lasting integration of research capacities existing across Europe.

Based on a particular scientific and technological theme and with a set of long term objectives, each Network aims to achieve a critical mass of competence and skills and to advance knowledge through the creation of a virtual centre of excellence. The networks will operate by fostering cooperation between existing centres of research excellence in universities, research centres, enterprises and technology organisations, expanding over time with the addition of new members.

The activities then undertaken by network members include not only joint programmes of research, but also associated collaborative activities, such as joint training activities, exchange of personnel, shared research infrastructures, equipment sharing and joint management of the knowledge produced. The minimum lifetime of these networks is five years, although it is anticipated that they will continue to function beyond the

period of EC funding, the amount of which is determined in relation to the value and capacities of the resources to be integrated by participants.

6.2.2 Integrated Projects

Sharing the same rationale as Networks of Excellence, Integrated Projects (IPs) aim to reinforce European competitiveness or to contribute to the solution of important societal problems through the mobilisation of a critical mass of research and technological development resources and skills existing in Europe.

Integrated projects will have clearly defined scientific and technological objectives, and may include research, technological development and demonstration activities. They are expected to cover innovation and the dissemination, transfer and exploitation of knowledge. They will incorporate a degree of flexibility, allowing participants greater autonomy to modify the joint programme, to add new partners and to launch new activities in response to changing circumstances.

To reflect the more ambitious nature of the projects, they will be funded for up to five years, with a corresponding allocation of resources, possibly reaching tens of millions of euros. Inter-related sub-projects may be integrated through a unified management structure.

6.2.3 Article 169

Article 169 of the EC treaty states: “In implementing the multinational framework programme the Community may make provision, in agreement with the Member States concerned, for participation in research and development programmes undertaken by several Member States, including participation in the structures created for the execution of those programmes.” There are no real experiences with this “géométrie variable” type of programmes yet. In big industries like aerospace there is a possibility that such initiatives will evolve. Also, a few member states, since the late 1990s, had begun to open some of their research funding programmes for international participation; e.g. a programme (PRO INNO) co-funding R&D cooperation of SMEs of the German Federal Ministry for Economic Affairs and Technology (see www.bmwi.de/Homepage/Politikfelder/Technologiepolitik). The level and extent of the member states' readiness to open their schemes and of the Community participation will not alt least depend on how difficult the related negotiation and decision procedures will be.

6.2.4 New, systemic instruments in national contexts

For about a decade now, we have been witnessing the emergence of new, systemically oriented RTD and innovation policy instruments in Europe, not at least reflecting the spreading of the "systems of innovation" approach of understanding the dynamics of research and innovation (OECD 1999). Such new, mostly rather experimental instruments may be characterised (see Smits/Kuhlmann 2002) as aiming at

(1) *management of interfaces*; i.e. striving for the building of bridges and stimulating the debate, not limited to bilateral contacts but also focuses on chains and networks at system level;

(2) *building and organising (innovation-) systems* by the facilitation of new combinations (Neue Kombinationen) and deconstruction (creative destruction) of systems, initiation and organisation of discourse and alignment, consensus;

(3) providing *a platform for learning and experimenting*, such as: learning by doing, learning by using, learning by interacting and learning at system level;

(4) providing an *infrastructure for strategic intelligence*, i.e. identifying sources (Technology Assessment, Foresight, Evaluation, Bench Marking), building links between sources, improving accessibility for all relevant actors (Clearing House) and encouraging development of the ability to produce strategic information tailored to the needs of the actors involved;

(5) *stimulating demand articulation*, strategy and vision development, facilitating the search for possible applications of future science and technology, developing instruments that support discourse, vision and strategy development.

An *example* representing several of the above characteristics is the *German "Futur" initiative*, a new kind of foresight process. Futur is run on behalf of the Federal Ministry of Education and Research (BMBF) as a means of priority-setting for future innovation-oriented research policies. Futur is oriented towards the identification and inclusion of societal needs in future research agendas. "Leading visions" (*Leitvisionen*) are supposed to be the major outcomes of the process which shall be translated into funded research programmes or projects. The participation of a broader audience in various kind of workshops and the combination of different communication and analytical methodologies are characteristics of the process.

Futur is intended to introduce "fresh ideas" into the research funding portfolio of the BMBF, by way of bypassing the traditional mechanisms for agenda-setting and prioritisation. The conventional process is characterised by a close and rather intransparent interaction between research institutions, industry, programme agencies (*Projektträger*) and ministerial bureaucrats in charge of research funding. The process started with *workshops* in early summer 2001 and is still running. Actors from industry, science, media and others were invited. These persons were identified because of their more broad, general knowledge. They were not supposed to be "specialists" in the narrow sense. From this first list, a kind of co-nomination process was conducted identifying about 1500 persons. Workshop participants were asked what they thought society might look like in the year 2020. As a next step, an "*open space conference*" was organised. The purpose of this conference was that "*focus groups*" should be founded and focus their themes. The groups had to produce "profiles" of their themes and a kind of competition was organised to write an interesting profile that meets a set of criteria given (new theme, societal need orientation, research link etc.). Out of the focus group's themes BMBF selected twelve areas to be more thoroughly debated in Futur. The Ministry organised an in-house workshop with department and division heads as well as the project management agencies' representatives asked to score the thematic areas according to their relevance. A similar process was organised in the internet: the persons already identified for the initial workshops were asked to give their votes. Finally, the Ministry asked the Futur participants to develop five of the twelve themes in detail, as a basis for the implementation of related funding measures. In the course of summer 2002 the Futur initiative will be evaluated by an international peer panel. Below we consider more broadly the evaluation of foresight as a policy instrument.

6.3 EVALUATION IN THE EUROPEAN RESEARCH AREA

Clearly, the precise form of specific future policy instruments at the time of implementation will determine the combination of evaluation methodologies that are relevant and their mode of implementation. At this stage it is more pertinent to consider the types of evaluation issues that the concept of the ERA and the instruments within it are likely to give rise to. This section considers these issues, as well as the specific evaluation issues associated with the two main future policy instruments. More general issues related to the ERA are also considered, such as the emphasis, within instruments, on personnel mobility, and on the relationship between benchmarking and evaluation.

6.3.1 Integration

The move towards greater integration of research activities, through both the networks of excellence and the integrated projects poses particular challenges for evaluation. By design, the whole greater is than the sum of parts, and so the ways in which the effects of integration can be assessed must be considered.

With the greater integration proposed in the ERA, evaluation at a European level must look beyond the evaluation of the Framework Programme, in terms of both scope and methodology. Closer integration of research policies requires mutual understanding of what has been achieved and how and of the distribution of benefits.

Developing greater linkages between national activities, must take into account the national variations that exist, such as in the institutional settings in which the type of work is carried out. For example, certain types of cancer research could be carried out in a university in one country, a branch of a national research organisation in another, a central laboratory in a third and by a non-governmental charitable foundation in a fourth. In each case, the core expertise has been acquired and developed through different sources, with different objectives, and with different modes of operation. An understanding of the dynamics of each national and institutional context maybe required for a thorough evaluation.

6.3.2 Excellence

The emphasis on research excellence requires consideration of the concept and measurement of “excellence” and efforts to move towards shared notions of both quality and excellence, and thus to the setting and adoption of standards. It could be argued that the development of a European Research Area requires a corresponding development of a “European Evaluation Area” in which there is a common methodological and procedural understanding that allows members to accept and validate each other’s findings. Thereby a proposal that a particular institution or centre meets a particular level of excellence in some dimension of its performance could be treated unequivocally.

6.3.3 Networks of Excellence

The long-term nature of Networks of Excellence and their mode of operation pose challenges for their evaluation. Research conducted by network participants should fall within the priority theme areas of the Framework Programme, or should respond to emerging policy needs, although the outputs and impacts of research may not be precisely specified at an early stage, to provide readily measurable and verifiable objectives.

Rather, it is anticipated that a long term evaluation perspective should be applied. The principal focus of evaluation of the operation of Networks of Excellence will centre on the added value they generate, bringing together the expertise of individual institutes into something larger than the sum of the parts, and its persistence beyond the initial period of funding.

The distinction between activities funded directly by the Commission and the other core activities and expertise of each of the network members is blurred, and thus a further challenge to the evaluator, as both must be taken into account. The true measure of the impact of the policy instrument will not be the productivity of a shared activity, as in a funded project, but in the enhancement to overall scientific productivity.

Any evaluation of the economic impacts of basic research must acknowledge the multiple dimensions of its effects on the economy, and, in this case, should pay particular attention to the production of human resources trained in the context both of scientific excellence and European added value.

Where concrete proposals for the evaluation of Networks of Excellence have been put forward these have centred on the arrangements for ex ante evaluation and monitoring. It is envisaged that the evaluation of project proposals will be based on a “peer review” approach, but may combine different methods depending on the characteristics of the areas under consideration, such as the use of independent expert panels, a referee system, and hearings of project representatives.

For the selection of the participants of these new instruments, it is indispensable to establish criteria for measuring excellence and quality. The long-term nature of Networks of Excellence requires a long-term evaluation perspective with a multi-dimensional assessment, focusing e.g. on the generated added value or the contribution to overall scientific productivity. Concrete proposals for Networks of Excellence can be evaluated in an "extended peer-review" approach, centring on eligibility criteria like excellence (number of publications or patents), managerial competence, Community added value, and the potential contribution to the integration of scientific efforts and the advancement of the scientific field; where necessary such extended information would have to be drawn from additional "background studies". Also, it is highly recommended that the peer panels undertake on-site visits of the "principal investigator", thus facilitating a critical debate between applicants and peers (see the Deutsche Forschungsgemeinschaft's (DFG) practice of appraising applications for multi-annual, multi-site integrated thematic research groups). An external scientific and technological monitoring council could monitor the progress of the network and assesses its work. The principal difficulty likely to be faced is that the size of individual grants limits the number which can be awarded. In these circumstances choices will have to be made between competing fields of research. The principles of peer review do not function in such circumstances, necessitating some higher strategic criteria for selection and a process to implement these. Ultimately this could be enhanced by information generated through foresight and other forms of strategic intelligence.

6.3.4 Integrated Projects

In terms of evaluation, the socio-economic dimension clearly has priority for this instrument. However, the approach is somewhat different to that for evaluation of smaller isolated projects. In the case of the Integrated Projects the project should achieve a critical scale whereby the strategic direction of a sub-sector of the economy is affected. This implies that the evaluation needs to engage with the socio-economic status of that sub-sector to understand the potential and actual impact of the IP (and the rationale for carrying out a project in this area). Effects could include such broader aspects as market structure. This can be described as a meso-level evaluation. Several of the approaches mentioned in earlier chapters of this book are applicable here. In terms of relevant past experience, the evaluation of projects such as JESSI (a successful joint EU/EUREKA programme in microelectronics) provide some relevant experience.

6.3.5 Mobility and Cooperation

More generally, the ERA objective of increasing mobility of researchers within and beyond the European Union, and in the context of different research policy instruments, raises broader issues for socio-economic evaluation. In particular, while the success of measures to promote researcher mobility may eventually be manifest in measurable impacts and outputs, studies to assess capability and potential are more relevant to

study the state and dynamic of the knowledge economy, and the role of science, particularly a high quality scientific workforce, in sustaining and attracting economic activity.

6.3.6 Benchmarking

One dimension of assessment activity of the European Research Area, which is clearly relevant to evaluation and where progress is currently being made, is the idea of benchmarking policies for science and innovation. With its origins in the industrial domain[78], benchmarking aims to identify and spread best practice. To do so, it relies on the same types of data as evaluation, such as bibliometrics and peer assessments, and draws on a similar vocabulary, with terms such as "relevant performance indicators", "qualitative understanding of best practice", "monitoring mechanisms" in common use. However, for evaluators, the use of benchmarking data should be placed on the context of the systems that generate them. For example, in the domain of economic evaluation, indicators such as patents and the income from intellectual property are highly context-dependent and should be used with due caution.

6.4 EVALUATION ISSUES AND METHODOLOGIES

6.4.1 Monitoring, analysing and assessing of changing national and sectorial innovation systems at National, Regional and Sectoral levels

For decades now, we have been witnessing in Europe a co-evolution of regional, national and European research and innovation systems and related policy arenas, the latter meanwhile merging into towards a multi-level, multi-actor systems (Kuhlmann, 2001; Georghiou, 2002). Regional, national and transnational levels undergo a re-distribution of tasks, thereby experiencing new functional and informational linkages, vertically and horizontally. Thereby, the integration and redistribution is proceeding with different speed across Europe: initiatives of the "géométrie variable" type have been suggested repeatedly and will be implemented (§ 169, 6th Framework Programme).

In an extrapolation of this development we would see regional or national authorities concentrating their efforts on the competitiveness of "local" innovation systems, while the EU Commission – instead of running cumbersome own funding programmes – would "mediate" between the competitors and "moderate" their conflicts, and would take care of horizontal policy coordination. Public investment in, and regulation of RTD and innovation would originate mainly from regional or national initiatives and sources – but it would be concerted and matched with any parallel activities throughout Europe. Here, an important task of the EU Commission would be to carefully evaluate and facilitate the transferability of funding instruments, developed in heterogeneous national, regional or sectoral contexts, across Europe, thereby providing a *forum* to debate the degree of immediate imitation versus the need of "domestication" (Silverstone & Haddon, 1996); consider e.g. – as a thought experiment – the degree of transferability vs. requested domestication of the German "Futur Process" for innovation-oriented research funding priority-setting. Such debates would have to be grounded on the results of related policy evaluations as well as other sources of Strategic Intelligence (see section 5 of this volume). The EU Commission would have to facilitate the production of related intelligence inputs, such as:

- a *new breed of "national impact studies"*. Since the early 1990s, studies with a strong focus on the impact of EU FPs on Member States' science and technology policy and national actors in the science and economic sphere have been carried out. Laredo (1990) examined the role of public and academic research institutions in the FPs, Georghiou et al. (1993), Reger and Kuhlmann (1995), and later a full series of similar studies (e.g. Luukkonen and Hälikkä, 2000) drew conclusions about the impact on national academic institutions as well as national industry and research organizations, thereby shedding light on the interaction between European and national policies. These national impact studies, though certainly milestones in terms of their methodological stance towards impact measurement, were, however, not full evaluations of the FP as they focused largely on the effectiveness of the programme's impacts. Nonetheless, they remain the most detailed cross-Framework examinations ever done since the introduction of European RTD policies. As a result of this experience, the Commission made attempts at designing a common core methodology for subsequent impact analyses. A new breed of "impact studies" would in particular take care of the interplay between national,

[78] See for example the document on CORDIS from the Commission entitled Comparing performance: a proposal to launch a benchmarking exercise on national R&D policies in Europe.

regional and EU RTD initiatives, combined with an analysis of the dynamics of sectoral innovation systems, not bothering about national borders.

- Consequently, in an integrating European research and innovation area the *study of sectoral innovation systems* will become an indispensable part of future RTD evaluation exercises.
- For the same reasons serious attempts would have to be made to extrapolate future research and innovation systems as well as policymaking mechanisms, as a basis for the design and ex ante evaluation of the next round of EU RTD policies ("7th Framework Programme") (see Kuhlmann 2001; Georghiou 2002).

6.4.2 Example: Evaluating Foresight

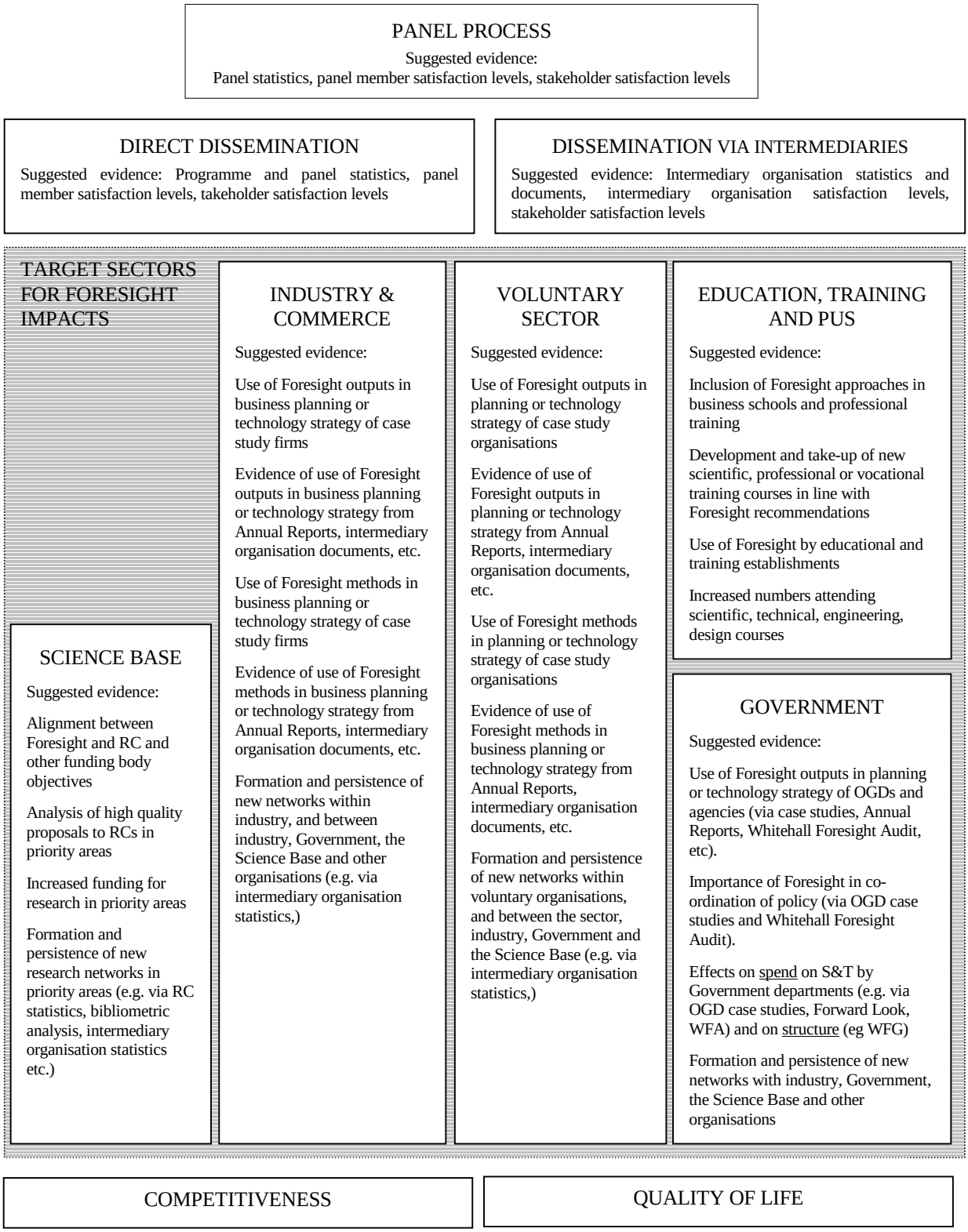
As noted above the German Futur is being evaluated by an international peer panel supported by surveys of participants and users. The broader issue of evaluation of foresight is one major challenge in the coming decade. Several countries are coming into a second cycle of foresight activity and are keen to learn systematically from their own and others' experiences. So far, most evaluation of foresight has been of a fairly ad hoc nature. As with many of the new policies under discussion, a key problem is that the range of stakeholders has been broadened to include social as well as scientific and economic actors.

The driver for foresight has been the provision of a shared space in which firms who have to innovate in concert with their customers, suppliers, academic collaborators and regulators can develop strategies which reduce uncertainty (Georghiou, 1996) or as martin and Johnston have put it, foresight is wiring-up the national innovation system (Martin and Johnston, 1999).

How then should foresight be evaluated? In a recent project in collaboration with Wiseguys Limited, PREST developed a framework for evaluation of the second UK programme. One of the key design considerations was to make allowance for the fact that the Programme relies to a great extent upon volunteers and upon the formation of networks. Both of these require a light touch so as not to disturb the effect being evaluated. One means adopted was to engage the panels which drive the programme in the development of both the framework for evaluation and the eventual measures which would be applied. This would build their commitment as stakeholders in the evaluation and help them to understand how it would benefit them. A further consideration was the importance of process issues – a great deal of foresight effort is involved with building a foresight culture and fostering particular ways of thinking. With a narrow base of expertise available and a typically two-year turnover among participants it was essential that the evaluation should provide some means of capturing, codifying and disseminating the knowledge base of the programme. For this and other reasons a real-time evaluation was recommended.

The hardest task was to reconcile the specific needs of foresight with broader government practice designed for more direct forms of business assistance and still founded upon a market failure rationale. The eventual compromise was to describe the Foresight Programme in terms of an indicator framework (see below) which covered the main dimensions of performance and intended impacts. The evaluation design team stressed the need to underpin all of these indicators with case studies which would facilitate interpretation of the data. In the event, this approach was superseded, at least in the short-term, by an internal high-level review prompted by the dissatisfaction of the responsible minister with the initial outputs of the Programme. The full evaluation is still pending.

Figure 11 Evaluation Framework for the UK Foresight Process



6.5 CONCLUSION

The flow of new policy instruments is likely to continue as policymakers attempt to adapt to changing circumstances and to innovate in the ways to stimulate innovation. Likely future directions include more effort to measure capacity (including the development of human capital where Bozeman and Gaughan (2000) are among those advancing the concepts for evaluation through their analyses of career developments using curricula vitae. Similar efforts are needed to assess the evolution of scientific infrastructure. Some preliminary ideas on how to benchmark these through equipment surveys are put forward by Georghiou, Halfpenny and Flanagan (2001).

A broader challenge for evaluation will be to assess the combined effects of different mixes of policies. This is a pressing issue as policymakers have come to realise that a portfolio of measures is likely to have greater effect than an imbalanced system which simply tries to reinforce one or two apparently successful measures.

6.6 COMMENTED BIBLIOGRAPHY

Selection of "national impact studies":

Georghiou L. et al. (1993): The Impact of European Community Policies for Research and Technological Development upon Science and Technology in the United Kingdom, London: HMSO

Larédo, P. (1990): L'impact de programmes communitaires sur le tissu scientifique et technique francais, Paris: La Documentation Francaise

Luukkonen, T., Hälikkä, S. (2000): Knowledge Creation and Knowledge Diffusion Networks. Impacts in Finland of the EU's Fourth Framework Programme for Research and Development. Publication of the Finnish Secretariat for EU R&D 1/2000. Painotalo Miktor, Helsinki

Reger G., Kuhlmann S. (1995): European Technology Policy in Germany. The Impact of European Community Policies upon Science and Technology in Germany, Heidelberg: Springer

Studies or papers on the future of EU RTD policies:

Georghiou, L. (2002): Evaluation of research and innovation policy in Europe – new policies, new frameworks? In: Shapira, Ph., Kuhlmann, S. (eds.): Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe, Cheltenham (E. Elgar) (forthcoming)

Kuhlmann, S. (2001a) 'Governance of innovation policy in Europe – Three scenarios', In: Research Policy, Special Issue „Innovation Policy in Europe and the US: New Policies in New Institutions", edited by Hans K. Klein, Stefan Kuhlmann, and Philip Shapira, vol. 30, issue 6/2001, 953-976

Information on "New Instruments"Evaluation Approaches and Policy Rationales:

Bozeman, B and Gaughan, M (2000) Evaluating Scientific and technical Human Capital: An Event History Approach, in Shapira, P and Kuhlmann, S (eds), Proceedings of the 2000 US-European Workshop on Learning from Science and Technology Policy Evaluation, Bad Herrenalb, http://cherry.iac.gatech.edu/e-value/

Georghiou, L. (1998) Issues in the Evaluation of Innovation and Technology Policy, Evaluation Vol 4(1): 37-51

Georghiou, L, Halfpenny, P and Flanagan, K, Benchmarking the provision of scientific equipment, Science and Public Policy, August 2001, pp303-311

Georghiou, L., Rigby, AJ and Cameron, HM, (eds) (2002) *Assessment of the Socio-economic Impact of the Framework Programme* available from http://www.cordis.lu/

OECD (1999): Managing National Innovation Systems, Paris (OECD)

Seminar on the New Instruments of the Framework Programme 2002-2006: networks of excellence and integrated projects - 20.04.01 Discussion paper: Networks of excellence

Seminar on the New Instruments of the Framework Programme 2002-2006: networks of excellence and integrated projects - 20.04.01 Discussion paper: Integrated projects

Silverstone R. & L. Haddon (1996): 'Design and the domestication of information and communication technologies: technical change and every day life', in: Silverstone, R. & E. Hirsch (eds.). (1992) 'Communciation by design. The politics of information and communication technologies', Oxford: Oxford university Press.

Smits, R., Kuhlmann, S. (2002): Strengthening interfaces in innovation systems : rationale, concepts and (new) instruments. Paper for workshop "New challenges and new responses for S&T policies in Europe", Brussels, April 22-23, 2002 (EU STRATA)