PLATTFORM ft eval
Forschungs- und Technologieevaluierung

THEMA:

New Policy Instruments, New Challenges for Evaluation:

Nr. 17
Mar. 2003

BUNDESMINISTERIUM FÜR BILDUNG, WISSENSCHAFT UND KULTUR,
Minoritenplatz 5, A -1014 Wien
Mag. Markus Pasterk
markus.pasterk@bmbwk.gv.at
Ursula Suda
ursula.suda@bmbwk.gv.at

BUNDESMINISTERIUM FÜR VERKEHR, INNOVATION UND TECHNOLOGIE;
Renngasse 5, A-1010 Wien
Dr. Rupert Pichler
rupert.pichler@bmv.gv.at

BUNDESMINISTERIUM FÜR WIRTSCHAFT UND ARBEIT;
Stubenring 1, A-1010 Wien
Dr. Wolfgang Neurath
Wolfgang.NEURATH@bmwa.gv.at

FFF-Forschungsförderungsfonds für die Gewerbliche Wirtschaft,
Kärntnerstraße 21-23, A-1015 Wien
Mag. Klaus Schnitzer
klaus.schnitzer@fff.co.at

FWF-Fonds zur Förderung der wissenschaftlichen Forschung,
Weyringergasse 35, A-1040 Wien
Dr. Rudolf Nowak
novak@fwf.ac.at

JOANNEUM RESEARCH, Institut für Technologie- und Regionalpolitik,
Wiedner Hauptstraße 76, A-1040 Wien
Mag. Wolfgang Polt wolfgang.polt@joanneum.at
Mag. Klaus Zinöcker klaus.zinoecker@joanneum.at

KMU Forschung Austria,
Gusshausstrasse 8, A-1040 Wien
Dr. Sonja Sheikh
s.sheikh@ifgh.ac.at

ARC Seibersdorf research,
2444 Seibersdorf
DI Anton Geyer
anton.geyer@arcs.ac.at

TECHNOPOLIS,
Prinz Eugen Straße 80/12, A-1040 Wien
DI Fritz Ohler
fritz.ohler@technopolis-group.com

TECHNOLOGIE IMPULSE GESELLSCHAFT TIG,
Grillparzerstr. 7/8, A-1010 Wien
Dr. Dorothea Sturn
dorothea.sturn@tig.or.at

ÖSTERREICHISCHES INSTITUT FÜR WIRTSCHAFTSFORSCHUNG,
A-1103 Wien, PF 91
Mag. Gernot Hutschenreiter
gernot.hutschenreiter@wifo.ac.at

WWTF - Wiener Wissenschafts-, Forschungs- und Technologiefonds,
1090 Wien, Währinger Straße 3 / 15a
Dr. Michael Stampfer
michael.stampfer@wwtf.at

Klaus Zinoecker

Preface

New policy instruments create new challenges for evaluation, both on a national but also on an European Level.

Birgit de Bossezon (DG Research) sees a new challenge for the Community research evaluation system and describes ongoing discussions, e.g. policy linkages and reflections within the Commission. According to her, it is intended that the first results of the work on evaluation concepts will be available in the first semester of 2003.

Competence Centre programmes have become highly popular in the last years in a large number of countries. The basic idea is to link science and industry via long-term common RTDI cooperations in publicly funded infrastructures. The evaluation of such centres, both ex ante and interim, is a challenge due to the complexity of these multi actor - multi measure ventures. In their article Michael Stampfer (WWTF) and Harald Hochreiter (TIG) describe the design of the mid-term evaluation process (Four Year Evaluation - 4YE) of the Austrian K plus Competence Centre programme and first experiences with the implementation. K plus can serve as an example for a typical "complex funding programme", which is generally in accordance with the new European policy developments and new FP 6 instruments, especially for evaluation issues.

Wolfgang Polt (Joanneum Research) discusses new developments in evaluation methods and strategies at the European level and gives a short review of recent projects (ASIF, EPUB) : ASIF gave a broad perspective on the rationale of public RTD policies; EPUB is focussed more on the methodological side, as it has brought together researchers from various fields and disciplines. Both projects stress the need for a coherent evaluation strategy on the European level.

John Rigby (PREST) takes then a closer look on the ASIF project. The specific aims of the study were to examine in detail the range of justifications for the public support of RTD programmes and to carry out a comprehensive investigation of the state of the art in the evaluation of the socio-economic impacts with a specific focus on drawing lessons for future evaluation, especially in the light of the decision taken by the Commission to adopt the new policy instruments of the European Research Area. Rigby's article reports the main conclusions of the ASIF Study on the justifications for RTD support and selection of appropriate policy instruments, furthermore on the key question of programme evaluation, which is how to understand best net programme impacts, what kinds of evaluation methods and techniques are most appropriate for the task of evaluating the framework programme and the challenges faced by the imminent introduction of the European Research Area concept.

In his article, Jaime Rojo examines more closely the socio-economic dimension of the evaluation of RTD policy foundations. He points out that evaluation plays a relevant role in the design and improvement of RTD policy by bringing accountability, transparency, and rationalization into the policy making process, but also, that the increase of multi-disciplinarity as the mode of organisation of RTD and the multiplicity of actors involved requires a re-examination of traditional evaluation approaches.
According to Rojo, the evaluation of the socio-economic impacts of science and technology policy is becoming an increasingly relevant topic in policy making. At present, most RTD policies define specific socio-economic goals against which achieved results should be evaluated. This is not an easy task given the uncertainty and intangible character of science and technology. At the same time, evaluation practices have to evolve in order to cope with the new extended rationales applied in science and technology policy. Despite the availability of a number of evaluation methods, there is scope and need to look for further methodological improvements in evaluation. At present, says Rojo, consistent evaluations can be conducted at the project level, but undoubtedly more thorough evaluations at programme or policy level will require advances in knowledge both in the causal relations between inputs and outputs and in the way of arriving to meaningful procedures for measuring and in aggregating these outputs.

In his book review, Spyros Arvanitis (ETH Zürich) discusses Andreas Fier's "Government Funded Industrial Research in Germany" (Staatliche Förderung industrieller Forschung in Deutschland"), "a well-conceptualized and well-structured study which I warmly recommend for reading for everyone interested in the economic effects of technology policy, not only in Germany" (Arvanitis).

Finally, let me draw your attention to an upcoming event in Vienna in May 2003. Various evaluation studies on Government funded R&D Activities have attracted increasing attention of researchers and policymakers in Europe. The ZEW (Germany), the JOANNEUM RESEARCH and the PLATTFORM (Austria) together with the the Federal Ministry for Education, Science and Culture (BMBWK) have organized a conference

- to present recent theoretical, qualitative and quantitative studies,
- to discuss scientific contributions to the evaluation of government funded R&D activities,
- to discuss experiences from US evaluation and standards (SBIR, GPRA aso.) for European R&D policies.

The conference is designed for two days and it is sponsored and participated by
(a) the National Research Council (NRC), USA
(b) the Austrian National Bank (OENB), Austria
(c) Organisation for Economic Co-operation and development (OECD), France.

You will find more details in this Newsletter and in the internet www.fteval.at/conference .

Klaus Zinöcker
Plattform Forschungs- und Technologieevaluierung GesbR
& Joanneum Research
Wiedner Hauptstraße 76, 1040 Wien
klaus.zinoecker@joanneum.at
www.fteval.at

PLATTFORM ft eval
Forschungs- und Technologieevaluierung

Birgit De Boissezon

A New Challenge for the Community Research Evaluation System

New strategy

A review into how Community research evaluation is organised and implemented has over the last months been the subject of preliminary reflection and discussions. The aim is to develop an updated strategy that meets the changing needs of policy makers in the wake of major new initiatives such as the European Research Area (ERA) and in that context, the launch of the 6th Framework Programme. In addition, strategy must take account of some of the wider political, economic and organisational changes that are taking place, in particular: enlargement of the European Union (EU); the Lisbon strategy for the EU to become the most competitive and dynamic knowledge based economy in the world; and the increased emphasis on governance and better lawmaking. Also to be considered is the growing body of knowledge on how to measure and evaluate the impacts of scientific and technological research. The strategy will attempt to set out the aims, actors, roles, tools, and implementation principles for Community research evaluation.

Current system

The current system for evaluation of Community-funded research has been in place since 1996 and is based on four components: an Annual Monitoring of programmes; a 5-Year Assessment of the Framework Programme (FP) and its Specific Programmes (SP); ad-hoc evaluation activities by the separate research programmes and ad-hoc impact studies at a national level. By and large this system has worked well and been commended in both external and internal assessments. But new policies require new tools and the major changes taking place in the research environment mean we should also examine new approaches to evaluation.

Changes to the research landscape

It is worth briefly reminding ourselves of what have been the major recent changes to the landscape for Community research. In January 2000 the European Commission proposed the creation of a European Research Area [1] . This project was endorsed at the Lisbon European Council on 23-24 March 2000 by the Heads of State and Government of EU countries and a series of objectives with an implementation timetable up to 2010 were set [2]. The concept was further developed in a Communication in October 2002 [3]. The European Research Area concept aims to create an 'internal market' in research, the restructuring of European research fabric, and the development of a European research policy. The 6th Framework Programme which was launched at the end of 2002 has been designed to contribute to integrating, structuring and strengthening the foundations of the European Research Area. ERA is extending the horizons of Community research interests and activities, forging a new relationship between Community and Member State research. Accordingly, Community research evaluation will have to find ways of building connections with the evaluation of research at Member State levels. The new instruments in the 6th Framework Programme, especially Networks of Excellence (NoE), Integrated Projects (IP) and ERA-Net are important in this respect.

Key questions

The requirement to undertake evaluation and monitoring is set out in the Decisions for the Framework Programme and the Specific Programmes [4] , in the Regulation laying down detailed rules for the implementation of certain provisions of the Financial Regulation [5] , and different Commission Communications [6] , including most recently the Communication on 'Better Lawmaking' [7] . In effect these requirements represent the baseline for any proposed changes to the system, but at the same time allow scope for new specific arrangements. Some of the practical issues being explored at present include the use of expert panels as against extended use of professional evaluators. With this issue also arises the balance between internal and externally run evaluation activities. How can the Commission reconcile the need for independent external advice against the need to maintain ownership of the evaluation process and through this, support learning? How are the different levels of evaluation to be brought together, say in the case of projects, programmes and policy. A specific concern is to design evaluations in a way which promotes their use and to develop better connections between the ex-ante assessment of programmes and ex-post evaluations. Finally, what about the tools for evaluations and especially the collection and use of data? How

can better use be made of the data that is collected already from projects by Commission services? Is there a need for better standardisation of approaches across the different types of research evaluations, or rather for best practice approaches.

Some concepts

When dealing with these questions it is important to keep in mind that whatever evaluation system is proposed should have considerable flexibility. It would be impossible and inappropriate to suggest that all research evaluation could be done in the same way, particularly so given the complexity of Community research policy and activities. Accordingly, one of the ideas being looked at is to create high level aims and principles for the evaluation system, reflecting a combination of overall policy needs and best practice in terms of how evaluation can be implemented. These would serve to better link evaluation with policy while at the same time promoting better alignment of evaluation activities and improved effectiveness. Under such an approach the aims of evaluation could, for example, be collectively described as evaluation reporting domains – in effect the broad categories of questions and issues against which evaluation activity could be expected to report. This could include items such as: direct and indirect impacts; efficiency and effectiveness of process; and more diverse issues such as leadership of the Community research system. In parallel, a set of implementation principles could include topics such as: the importance of a systems approach to evaluation; the importance of learning about success; the need for strong ex-ante assessment; using evaluation to promote well-run projects; sound management of data; enhancing ownership and utility of evaluation activity and results; and the role of learning in evaluation.

This ongoing reflection has made extensive use of a range of different information sources, including the EPUB [8] and ASIF [9] reports from 2002 as well as formal and informal contacts with experts. It is intended that the first results of the work will be available in the first semester of 2003.

Author: Birgit De Boissezon
Head of Unit "Planning, Programming, Evaluation"
European Commission
Research Directorate-General
Office: SDME 02/37
B-1049 Brussels

The opinions expressed by the author do not necessarily reflect those of the European Commission.

1 Commission Communications: Towards a European Research Area, 18.1.2000, COM (2000) 6 final; Making a Reality of the European Research Area: Guidelines for EU Research Activities, 4.10.2000, COM (2000) 612 final

2 'The Union has today set itself a new strategic goal for the next decade: to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion.' (Point 5 of the Lisbon conclusions, 24.03.00, SI(2000) 300

3 Commission Communication, The European Research Area: Providing New Momentum, 16.10.2002, COM (2002) 565 final

4 These specify that 'the Commission shall continually and systematically monitor, with the help of independent qualified experts, the implementation of the sixth framework programme' and ' before submitting its proposal for the next framework programme, the Commission shall have an external assessment carried out by independent highly qualified experts of the implementation and achievements of Community activities during the five years preceding that assessment'. OJ L 232, 29/08/2002, pp. 1 – 33. In addition, there is a requirement that in 2004 an evaluation will be undertaken by independent experts of the new instruments (integrated projects and networks of excellence) in the execution of the framework programme. (OJ L 294, 29/08/2002, pp.1– 40)

5 Article 1 of Regulation (EC) No 1687/2001 of 21 August 2001, published in OJ L 228, 24.8.2001

6 Focus on Results: Strengthening Evaluation of Commission Activities, 26.07.00, SEC (2000)1051

7 Commission Communication on Impact Assessment, 5.6.2002, COM(2002) 276 final

8 European Commission STRATA project HPV 1 CT 1999-00005, RTD-Evaluation Toolbox: Assessing the Socio-Economic Impact of RTD Policies, August 2002, EUR-20382-EN, available at http://epub.jrc.es/docs/EUR-20382-EN.pdf

9 Assessment of the Socio-Economic Impacts of the Framework Programme, April 2002, available at http://www.cordis.lu/fp5/monitoring/studies.htm

PLATTFORM ft eval
Forschungs- und Technologieevaluierung

Harald Hochreiter,
Michael Stampfer

Mid-Term Evaluations of the Austrian Competence Centre Programme K plus

Abstract

Competence Centre programmes have become highly popular in the last years over a large number of countries. The basic idea is to link science and industry via long-term common RTDI cooperations in publicly funded infrastructures. The evaluation of such centres, both ex ante and interim, is a challenge due to the complexity of these multi actor – multi measure ventures. This contribution describes the design of the mid-term evaluation process (Four Year Evaluation – 4YE) of the Austrian K plus Competence Centre programme and the first experiences with it.

The K plus Competence Centre Programme: Overview

The K plus programme funds collaborative research structures with a specified life time, set up to carry out top quality, long term and internationally competitive research and technological development (RTD) projects at a pre-competitive stage. The goal is to perform research that is important for both the academic world and industry and to develop human capital in areas which are relevant for a number of sectors / companies.

K plus Centres are closely linked to Austrian universities and/or other research institutions. Nevertheless they are incorporated as limited liability companies with co-operative governance structures. They get K plus funding for a seven year time span. A minimum of five industrial partners contribute a minimum of 40% to the total budget. The overall annual budget of a K plus Centre amounts to 2.5 – 5 million€ . Work is planned on a mid-term basis (2-5 years) along a commonly defined work programme. This programme consists of a number of areas and projects. Between 30 and 60 researchers and technicians are working in a centre.

Currently 18 K plus Centres are in operation, five of them have already undergone their mid-term evaluation ("Four year evaluation", 4YE). During their seven-year-term these 18 centres will have a total turnover of roughly 400 million € and employ a total workforce equalling 900 full-time equivalents.

Rules and guidelines of this strategic funding programme were designed by the Federal Ministry for Transport, Innovation and Technology (BMVIT). Technologie Impulse Gesellschaft, a federal funding agency, acts as programme manager for K plus. This implies contract negotiations, steering, funding and evaluating the centres.

Competence centre funding programmes have become popular over the last ten years in a number of countries. The overall aim is always to better link science and industry and to stimulate strategic industrial research. Prominent examples for such centre programmes are the NSF Engineering Research Centres (ERC) in the U.S., VINNOVA`s Competence Centres in Sweden or the Australian Cooperative Research Centres (CRC). All these programmes already have undergone centre evaluation rounds.

K plus evaluation procedures

The incorporation of international best practices into the K plus evaluation procedures was an important issue on the agenda from the start of the programme on. The design of different elements and stages of the evaluation procedures always involve(d) evaluation experts from Austria and abroad. This preparatory work resulted in elaborate procedures intended to ensure highest possible quality during the selection of centres as well as providing suitable instruments to enable both centre and programme management to steer the centres. The following lines give a short overview on the K plus evaluation procedures:

- Step 1: Ex ante evaluation of proposals: The first core element of the selection procedure for K plus Centres is a combined procedure of scientific-technical peer review and a thorough examination by economic experts. Six foreign peers from the scientific community with a strong background in industry-oriented research are selected by the Austrian Science Fund (FWF). They provide a detailed assessment of scientific-technological expertise in the consortium applying for funding, while another Austrian institution, the ERP Fund, checks the organisational aspects of the planned centre and the embeddedness in the industrial and regional structures. The success rate in this two stage procedure (outline and full proposal), cumulated over three selection rounds so far, is less than 50%.
- Step 2: Two Year Review (2YR): Some centres faced a short review after two years as additional requirement

when being approved for funding, in order to check the progress of the centre and to give room for first recommendations (to be investigated more thoroughly during the Four Year Evaluation). The 2YR involves one of the scientific peers from the ex ante evaluation, one expert from the FWF and the ERP-Fund as well as TIG and the regional co-funding organisation.

- Step 3: Four Year Evaluation (4YE): The second core element regarding evaluations is the Four Year Evaluation: Design, procedures and outcomes will be described in detail below.
- Step 4: Ex post centre evaluation after seven years: An evaluation of outcomes and impacts will be performed on centre level after K plus funding ends. Work on the design of these evaluations will be done in 2004.
- Step 5: Programme Evaluation commissioned by the Ministry or the Austrian Research Council. This final step will be an evaluation of K plus on the programme level. The decision when and how to perform this evaluation is not within the responsibility of the funding agency. Nevertheless it is important to state that the 4YE is structured in a way allowing a programme evaluator to draw upon ample and structured information.
- Ongoing activity 1:
  Monitoring of additionality. This activity is dealing with additionality on the level of the participating industrial partners. TIG has started in accordance with the ministry a monitoring survey that measu res relevant research and innovation data on the level of the participating firms in the first, fourth and seventh year of their K plus participation. Therefore it is possible to get an idea about input, output and behavioural additionality of these firms, though it is highly difficult to separate the effects of K plus participation from other (RTD) activities of those firms. To install a control group, this survey is linked to the Austrian Community Innovation Survey questionnaire. Note that due to synchronisation and attribution problems this additionality monitoring is not part of the formal funding decisions based on the 4YE – and that generally this instrument has to be interpreted cautiously.
- Ongoing activity 2:
  Learning mechanisms. Parallel to these steps the programme includes continuous approaches towards learning routines and application of Best Practises: Centre workshops (e.g. on evaluations) and international benchmarking activities are important in this context.

Four Year Evaluation (4YE): Key Facts

The main goal of the K plus programme is to improve science-industry cooperation by building mid- and long-term RTD structures based on existing competence. A commonly defined work programme is the core of these co-operative centres. The Four Year Evaluation therefore shall bring evidence on results, plans and integration steps. This important evaluation exercise is performed in the first quarter of the fourth year to give enough time for adapting to conditions and recommendations before the second funding period starts.

Given the complex structure, the long negotiation procedures and the mid-term orientation of those centres, there is limited scope for applying K.O.- criteria. Building up a competence centre takes time. Therefore only two explicit K.O.-criteria are included: severe mismanagement and lack of suitable RTD planning procedures. Contrary to ex ante evaluation the "burden of proof" during the 4YE is on the evaluators' side (not on centres' side). Nevertheless the outcome of the evaluation can result in substantial adaptations of the centre structure or the programme – as a worst case even the closing down of a centre is possible.

This mandatory 4YE procedure is set in the funding guidelines and is a kind of counterweight to the centres' planning and management autonomy after approval. The design of the 4YE includes ex ante and ex post elements: measurement of performance during the first three years is based on a set of indicators (see below), the plans for the years 5 to 7 (and beyond!) are appraised in view of the long-term strategy of the centre. The outcome of the 4YE is decisive for the second funding period (years 5 to 7): This means stop or go, amount of money granted and possible additional requirements.

The evaluation is performed by a mixed group of evaluators, the key elements are an analysis of the written information and a two-day site visit. The evaluation is commissioned by TIG, the formal decision on funding the second period is with the Federal Ministry (BMVIT). For details, see below.

Programme Goals – important building blocks for 4YE

For the 4YE the goals of the K plus programme apply according to the funding guidelines. The indicators and the questionnaires for evaluators and centres are based on this set of programme goals, which are of course abstract but nevertheless useful for the evaluation task. On this abstract level the goals are mandatory for the centres via funding guidelines and contracts, again not in a strict way, but as one overall set of benchmarks for all K plus Centres. Of course not all goals are of equal importance as some of them – like "raising of acceptance for RTD in public" – is harder to operationalise and to achieve than others like "build up of critical masses".

The overall strategic goal of the programme and therefore relevant for the centres and the 4YE process is "Improve Science - Industry Co-operation". All the following operational goals (see Box) are relevant for the 4YE process, though – as already mentioned – only severe mismanagement and planning defaults can lead to a premature close-down due to negative evaluation results:

4YE Programme Goals

- A: Build up and utilisation of knowledge and competences; Perform longer term, strategic RTD
- B: Secure / raise attractiveness of Austria (of region) as a high tech location
- C: Build up of critical masses
- D: Raise acceptance for RTD in public
- E: Participation in international RTD programmes
- F: Increase additionality
- G: Raise professional research management standards
- Z: Specific centre goals

Goals set by centres themselves come in as additional information. As the consortia bidding for K plus centres shall create long term RTD infrastructures boosted by considerable public funding, they set additional goals of their own. These centre specific goals are written down in the centre agreements. They are only subject to the 4YE as far as they can be operationalised and as far centres accept the evaluation of these goals. Some of the K plus centres have been proven to be more precise than others.

Relevant indicators for the 4YE

This set of evaluation indicators was published in 1999, soon after the first centres had their start. It is closely linked to the programme goals and to the ex ante evaluation. Some indicators are of a quantitative nature, some are mixed, a considerable number is purely qualitative. It was important to have two sub-sets, one counting, interpreting and valuing the direct results, and one measuring and judging the network and indirect effects. There are seven indicators regarding direct results of RTD performed in a centre: 1) Performance of planning and management, 2) publications, 3) patents, 4) successful tendering for RTD funds, 5) invitations, scientific prestige, 6) PhDs, young scientists, 7) conferences, workshops, visiting scientists. This subset is complemented by eight indicators regarding effects and outcomes of RTD with the centre and partners: 1) Scientific relevance, 2) long term co-operation, 3) strengthening of qualitative and quantitative RTD efforts in centres' industrial partner structure, 4) use of results on partner firm level, 5) new co-operation partners of the Centre, 6) co-operations between partners, 7) transfer of personnel and 8) general management and cultural issues.

All these indicators are linked to one or more goals and to questions in the evaluators' questionnaire. They are measurable in different ways. The expertise and experience of the evaluators is regarded as a core asset. It is more important to get a good qualitative judgement than a box full of non-interpreted quantitative indicators. So in many cases the ultimate "quality" question is: "Would this be a successful centre in your country?" . In this way there is a comparison to Australian, American, German, Swedish etc. standards. The centres know in advance which data they have to collect, there is also a strong link to the K plus reporting system.

The linking of goals, indicators and procedures was designed by TIG together with Austrian evaluation specialists.

Schedule for a centres' 4YE

The most important information base for evaluators is the so called "core document" written by the centre and provided in advance. The structure of the document is predefined and consists of two parts. The first part is a report on the years 1 to 3 with the main question: how did the centre perform and which results / outputs have been produced so far. The second part is the planning document for the years 5 to 7 which is evaluated on the background of the strategic outlook for the years 8+. Of specific importance is a presentation of the long term vision for the centre. The core document should not surpass 100 pages and its structure reflects the goals, indicators and questions to the evaluation group. Additionally the members of the evaluation group receive an information package about the K plus programme, the 4YE in general and about the specific centre.

Members of evaluation group

For TIG as the responsible funding agency it was important to constitute mixed groups to include the different perspectives regarded as necessary. This means: experiences from the ex ante process and people with a fresh view; experts for the individual centre and "generic" evaluators visiting all centres. So each centre is visited by an evaluation group including two different sub-groups. The first consists of three foreign scientific peers. They are individually chosen, one from ex ante evaluation group plus two new experts. The second sub-group is the so called standing group, also including three experts who visit all centres: As the chairman of the whole evaluation process could be won a high level foreign scientist with ample experience in RTD management. This position is filled by Prof. Haim Harari from the Weizmann Institute. Second is a "centre manager", a former director of a similar competence centre from Sweden, Sten Ljungström. Erik Arnold from Technopolis comes in as an international evaluation expert.
Apart from the six members of the evaluation group the following persons take part in the site visits. Having no vote, they act as information providers and observers: The ERP Fund for industrial expertise (1 employee; author of short expertise on industrial impacts), Austrian Science Fund (FWF, 1 member of scientific board, 1 employee), one representative of the co-financing regional government and TIG as secretariat, the latter providing information on the centres' financial performance. The minutes of the 4YE and all recommendations are decided upon and written by 3+3 voting members only.

Standing group: Principles

The standing group is core of the 4YE process and acts according to the following principles: Continuity: One core group visits all Centers. Governance and structure approach: No evaluation of scientific outcomes (which is covered by scientific peers), but on structural, organisational and policy questions. Management approach: Evaluation of management procedures and practises. Learning approach: This means seven year perspective; this includes formative elements. Internationally competitive and appraisal of special situation: This includes the 'can they make it anywhere' and 'do they make most of their potential?' questions. Finally it is about clear recommendations (together with peers): This means a clear yes/no recommendation to the ministry including maximum funding and conditions.

Site visit: Centre participation and agenda

During the site visit the evaluation group meets the centre director, key- and senior-researchers as well as junior researchers and representatives of industrial partners, for some questions also representatives of owners and important stakeholders of the centre. For the assessment their vision and their long-term strategy for and with the centre plays a vital role during the 4YE.
The agenda consists of a general presentation of the centre, followed by presentations of selected projects from the centres' research areas. There is time for discussing scientific-technical matters as well as organisational and management related topics (financial, legal, marketing, HR) along the structures of the core document which is also the structure of the evaluators' questionnaire. In most cases there is an on-site visit of laboratories and research facilities which also gives the possibility to discuss and talk to junior researchers not presenting their projects.

4YE Character of recommendations

The first "deliverable" of the 4YE is a short summary signed by all members of the evaluation committee after the closed session at the end of the centre visit. It contains the most important elements: Should the centre be funded for another three years (yes/no)? If yes, are there conditions and/or recommendations? Which amount of funding is appropriate?
A more extensive report is the drawn up by a member of the standing group ("Rapporteur") usually within two weeks after the 4YE. It provides comments on all sets of questions from the questionnaire and also works as a feedback to the centres. They receive this final report after the ministry took the funding decision (according to the recommendations).
Based on the decision of the ministry and on the findings included in the final report, TIG as the funding agency gets a negotiating mandate to implement all additional requirements and recommendations regarding adaptions concerning organisational structures, IPR, financial aspects and – in most cases – strategic orientation of the centres.

Lessons from the first 4YE's

At the time of writing (December 2002) five centres have undergone their 4YE. They all passed and will receive funding for the years 5 to 7. Yet all except one centre got their approval only under strict conditions and recommendations mainly concerned with IPR and – even more important – the long-term plans of the centres. Especially the point to come up with more explicit strategic planning mechanisms and documents was highlighted. On the other hand scientific-technical aspects were less frequently a matter of concern and so was the commitment by industry. This is a clear indication that the ex ante evaluation procedures are working well as these scientific and industrial issues are topics under close scrutiny during the selection process.
The concept of the 4YE seems to work well. The composition of the evaluation team secures a broad approach – and while the scientific experts bring in their expertise specific to each centre under review, the standing group ensures continuity, comparability and learning effects within the evaluation exercise. The fact that differences between centre structures can be adequately met is reassuring, so are the productive discussions and the unanimous decisions among the members of the committee after each of the evaluations.
Regarding the design and methodology it is important not to over specify such evaluations. This has two aspects: First such a competence centre is above all a social venture of a group of institutions and individuals coming from different worlds. This is a strong argument for qualitative, centre specific and learning approaches which have to be balanced against the evaluation-specific needs to get comparable figures, standardized procedures and managerial options. The second aspect is the balance between getting answered all interesting questions and the danger of getting thousand details instead of a clear picture. Here TIG chose the approach to structure in detail the written information and questionnaires, while leaving it to the evaluation group where to put the emphasis in the individual evaluations.
What lessons can be learnt for K plus and other programmes? Judging from the most frequently stated concerns there seems to be two topics. One is the question of allocation of intellectual property rights from jointly operated research programmes. The clear hint from all 4YE procedures was that the position of the centres had to be strengthened relative to industry. When negotiating the terms for funding in the years 5 to 7 this is an important part of the funding agencies' job: making sure that industries need for protecting results does not inhibit the centre to build a second area of business termed "Non K plus". Only if the centre manages to diversify into additional activities such as contract research for (additional) industry partners, participation in international projects (esp. EU framework programmes) and keeping their competence at an internationally competitive level will they be able to "live a life after K plus funding".
This leads us to the second lesson where the answer to the question is less clear. K plus centres receive significant amounts of funding for seven years. And as long as they continue to work in the pre-competitive field they are – and should be – entitled to a maximum of 60% public funding. But this 'guaranteed' funding stops at the end of year seven. Whatever their plans for continuation are, centres face the fact of a rather abrupt end in public funding – provided no other substantial source of income can replace K plus money. On the small scale the answer is easy: it is the responsibility of management and the owners of the centres to make sure that the transition from year 7 to 8 goes on as smooth as possible. This includes planning various scenarios and fine-tuning towards the end of the funding period. On a larger – and more political – scale there will have to be an answer to the question whether federal and / or regional government should offer a longer term financial perspective to keep at least the most promising of these cooperative innovation structures alive. Still this does not free the owners from their responsibility to have a clear vision why those centres should continue after K plus – but it might be one important element in the mix of income that allows to continue doing (pre-competitive) research instead of development or consulting work.

Lessons for upcoming European Research Area?

The European Research Area (ERA) is an answer to the challenges for Europe, given the increasing importance of RTD performance for global competition. The core of ERA will still be the main instrument of community RTD policy, i.e. the (6th) Framework Programme. In this case K plus can serve as an example for a typical "complex funding programme", which is generally in line with the new European policy developments and new FP 6 instruments. This holds also true for evaluation issues. Multi-factor ex ante evaluation is key for competence centres – similar to Networks of Excellence and integrated projects, where instruments shall be evaluated in form of an "extended peer review". The 4YE as mid term review is structured along a number of core indicators, most of them of qualitative nature. It is important not to fall into a kind of virtual quantitative exactness, but to allow experts to judge along meaningful indicators and categories – and to strike a balance between necessary control mechanisms and the management autonomy of complex consortia. Learning is one keyword, complementing but not replacing necessary control mechanisms when public money is in the game.
K plus is also a programme open for participation of foreign industrial and scientific partners. While this is a signal to

reflect the growing RTD links within Europe, the centres have to show in the evaluation that they themselves are able to link up with foreign partners, namely in international programmes. The 2002 FP 6 Expression of Interest round was very promising in this respect.
A difficult question in the K plus evaluation, but also for all kinds of multi actor instruments is the measurement of output additionality, namely in partner companies; behavioural issues are also important, but difficult to follow. In this case the K plus long term additionality monitoring of output and changes in partner firms parallel to the 4YE could be an interesting model for other programmes and initiatives.
Finally K plus has international evaluation boards, English as common language, and uses international best practice approaches. Transnational policy learning takes place on a bilateral level but also in a EU STRATA Thematic Network funded with FP 5 money. This network, co-ordinated by TIG, is called MAP and is about the management of complex 'multi actor – multi measure' funding programmes with a strong emphasis on selection and evaluation mechanisms. It includes a number of funding programmes from a broad range of countries (see www.map-network.net). Generally spoken there is a lot of effort to design and manage K plus evaluation procedures in a way that it can be seen as European good practice.

References:

Australian Government Publishing Service, 1995: Changing Research Culture - Australia 1995. Report of the CRC Program Evaluation Steering Committee
Georghiou, L.; Roessner, D., 2000: Evaluating technology programs: tools and methods, in: Research Policy 29 (2000), 657 ff
Mercer, D., Stocker, J., 1998: Review of Greater Commercialisation and Self Funding in the Cooperative Research Centres Programme
NUTEK, Sweden, 1997: The NUTEK Competence Centre Programme. First International Evaluation, Groups 1+2+3
Ohler, F., Stampfer, M., 1999: Konzept zur Evaluierung der K plus-Kompetenzzentren (Mimeo)
Parker, L., 1997: The Engineering Research Centers (ERC) Program - An Assessment of Benefits and Outcomes [www.nsf.gov/pubs/1998/nsf9840/nsf9840]
Petit, S., Vigneron, V., 1999: Bewerten der Folgen von Technologietransfers zwischen staatlichen Forschungslaboratorien und Privatunternehmen, in IPTS-Report (40), 42 ff
Roessner, J.D., Melkers, J., 1995: Evaluation of National Research and Technology Policy Programs in the United States and in Canada, in: Kuhlmann, S., Holland, D., 1995: Evaluation von Technologiepolitik in Deutschland. Konzepte, Anwendung, Perspektiven, 243 ff
Technopolis Ltd., Cross Border Cooperation within National RTD Programmes, 1999, Study for the European Commission
Zinöcker, K., Stampfer, M. et. al., Ein Evaluierungskonzept für das Kompetenz-zentrenprogramm K plus. Zielkataloge, Kriteriensätze, Abläufe, Materialien, 2001

Author: Harald Hochreiter joined the Austrian Ministry for Science and Technology during the design phase of K plus in 1998 and was subsequently Programme Manager for K plus in Technologie Impulse Gesellschaft. Since 2001 Programme Director K plus and in this function secretary of the 4YE Evaluation Committee.
harald.hochreiter@tig.or.at
http://www.tig.or.at

Author: Michael Stampfer, director of the new Vienna Science and Technology Fund, from 1992 – 1998 in the Austrian Ministry for Science and Technology. 1997 responsible for the development of the K plus programme. 1998 – 2001 Programme Director K plus in Technologie Impulse Gesellschaft.
michael.stampfer@wwtf.at
http://www.wwtf.at

PLATTFORM ft eval
Forschungs- und Technologieevaluierung

Wolfgang Polt

New Developments in Evaluation Methods and Strategy at the European Level – A Short Review of Recent Projects (ASIF, EPUB)

Background

Recent years have seen a renewed interest in further developing evaluation methodologies in the European context, especially with respect to the Framework Programme (FP) for research and technological development. The background of this development was on the one hand the increasing complexity of the evaluation task: each successive framework programme added to the number of goals including broader societal and environmental goals, instruments used within the FP proliferated to span from joint pre-competitive R&D projects to diffusion-oriented measures. On the other hand, there was a perceived need for greater accountability and proof of the effects of the FP than could be supplied by the monitoring and assessment methods in place.

Against this background, the commission initiated some efforts in research into evaluation methodologies, namely the project ASIF ('Assessing the Socio-Economic Impacts of the Framework Programme') and the thematic network EPUB ('Socio-Economic Evaluation of Public RTD Policies').
The Institute for Technology and Regional Policy of Joanneum Research participated in ASIF (which was co-ordinated by PREST/ University of Manchester and co-ordinated EPUB jointly with the Institute for Prospective Technology Studies/Seville.

Outputs

The reports (or CDs) of these projects are available from the respective web-sites
http://les.man.ac.uk/PREST/Publications/ASIF_report.html,
http://epub.jrc.es/evaluationtoolbox/start.swf or on request from John Rigby john.rigby@man.ac.uk and Klaus Zinöcker Klaus.zinoecker@joanneum.at.
Books containing further elaborated papers on specific aspects and methods are forthcoming in Spring 2003 by Edward Elgar (http://www.e-elgar.com). These books will also be reviewed in this newsletter.

Content

ASIF tried to provide a broad perspective on the rationale of public RTD policies and had four major parts: (i) it looked into the justification of the role of RTD policies based on a review of recent literature, (ii) it examined the evaluation practice of the FPs and other relevant RTD activities with the aim of delivering lessons and suitable practical principles for evaluations, (iii) it provided case-studies of specific types of R&D support and evaluations (support for business R&D, support for the service sector, evaluation on the programme level, evaluation in social sciences, a case study on accession countries), (iv) ventured to propose future avenues for the evaluation strategy in the light of recent developments of the European research system.

While the review of the current practices provides ground for learning from 'good practice examples', the chapter on future direction for evaluation strategies underlined among others the increased necessity for more elaborate ex-ante evaluation procedures, especially when dealing with the large scale projects to be initiated with the new instruments of the 'Integrated projects (IP)' and the 'networks of excellence (NoE)'.

EPUB, while also taking into account the broader background, focussed more on the methodological side. As it brought together researchers from various fields and disciplines (economists, sociologists, political scientists, ..) it was able to cover a wide range of methods, of which the state-of-the-art was reviewed, and potentials for application to the specific policies measures the EU currently employs in the field of RTD policy were highlighted.

These methods included:
- Innovation surveys
- Econometric models on the macro and meso level (aggregate national or branch level analysis)
- Microeconometric approaches (covering firm level data)
- Control group approaches
- Cost-benefit analysis
- Expert panels and peer review
- Field studies and case studies
- Network analysis
- Benchmarking, foresight and technology assessment as complementing tools in the context of a 'system of distributed intelligence for S&T policy making'

Apart from the findings in the specific fields of methods,

which pointed to new areas of application in several corners of the FP and other EU RTD policy instruments, the general findings were, that though there is scope for improving the evaluation procedures currently employed, there is also need (a) to further develop methodologies (e.g. in the promising field of microeconometric methods), (b) to provide more timely and accurate data, which could be done best, if evaluative questions are already taken into account when designing the proposal assessment, the monitoring and reporting procedures for the individual projects.

Both projects stressed the need for a coherent evaluation strategy on the European level.

While addressed to European policy makers, both projects contain general considerations on evaluation methods and strategies alongside a wealth of information on specific evaluation examples from various member states and thus are also a source for 'policy learning' for policy makers and programme managers in individual member states as well.

Author: Wolfgang Polt
Joanneum Research
Forschungsgesellschaft mbH
Institut für Technologie- und Regionalpolitik – InTeReg
Wiedner Hauptstrasse 76
A- 1040 Wien
tel.: +43 1 581 7520 – 2822
fax: +43 1 581 7520 – 2820
wolfgang.polt@joanneum.at
http://www.joanneum.at/rtg

PLATTFORM ft eval
Forschungs- und Technologieevaluierung

John Rigby

ASIF – Evaluating Socio-Economic Impact

Background

In 1999, the European Commission DG Research supported a major study led by PREST (Policy Research in Engineering Science and Technology) of the University of Manchester to review the rationales for RTD programmes and examine the range of techniques with which their socio-economic impacts could be most effectively evaluated. The specific aims of the study were to examine in detail the range of justifications for the public support of RTD programmes, and to carry out a comprehensive investigation of the state of the art in the evaluation of the socio-economic impacts, with a view to drawing lessons for future evaluation, especially in the light of the decision taken by the Commission to adopt the new policy instruments of the European Research Area. This short article reports the main conclusions of the ASIF Study on the justifications for RTD support and selection of appropriate policy instruments; on the key question of programme evaluation, which is how best to understand net programme impacts; what kinds of evaluation methods and techniques are most appropriate for the task of evaluating the framework programmes; and the challenges faced by the imminent introduction of the European Research Area concept.

Justifications for Support and the Selection of Policy Measures and Instruments

Any RTD development activity needs assessment of the scale and depth of the socioeconomic impacts it generates. But RTD programmes also need review and testing to warrant that their rationales are valid. Evaluation is the activity which carries out both of these central functions of programme design and management; and while impact assessment often appears to be the more important of these two activities, evaluators must be fully aware of the rationales and principles which stand behind and underpin the programmes so that they recognize impacts and make sensible and coherent suggestions for their improvement.

Our knowledge of why and how government should fund RTD develops continuously through the work of evaluators, historians and economists studying the innovation system and the role played within it by government. In recent years, as the ASIF reports points out, there has been a mounting awareness that economic growth depends upon the structure of the innovation system, the capabilities of the actors within it and the growth, and distribution of knowledge. A key concept is now thought to be that of "failure" within different parts of the economic system. This usage of the concept derives from the work of evolutionary and structuralist economists and it has been applied to a range of features of the innovation process, such as learning, systems, selection, and knowledge processing. A concept rather more widely applied than the concept of "market failure" from neoclassical economics therefore, the structuralist concept of failure demands of RTD policy makers that they seek to preserve freedom of action, encourage variety and ensure that the capabilities of actors within the innovation process are continually enhanced, rather than narrowed by other features and forces within the economic process.

However, policy makers should not abandon their neoclassical textbooks: the modern approach, while leaning far more towards the structuralist perspective, involves a number of paradigms. Furthermore, the policy maker needs to shape programmes to take account of what we know about innovation at different levels of the economy: at the level of the economy as a whole, within sectors, and also within the firm itself. At the level of the firm, the resource-based view has become highly influential, entailing support at the level of different competences instead of that based on traditional organisational analysis.

Evaluating Programme Impacts – the Importance of Net Effects

The main part of the ASIF Report focuses upon impacts and the difficulties thrown up by the need to address the key issue of net impacts. When programme evaluators assess the impact of programmes, they seek to determine not only whether the programme made an impact, but also to assess the level of justification for the use of government money on intervention. Such differences between what events occurred with the programme intervention compared to what would have happened without it are the net effect of the programme. Net effect is often also referred to as additionality, but the concept of net effect or

PLATTFORM fteval
Forschungs- und Technologieevaluierung

Evaluation of Government Funded R&D Activities

www.fteval.at/conference

Recently, the impact of R&D policy instruments on innovation, structural change and technology leadership in European Economies is being widely discussed. Various evaluation studies have attracted increasing attention of researchers and policy makers. At the same time, the evaluation methods used to analyse the impacts of R&D policies have been improved and become more sophisticated.
The aim of this conference is to discuss recent scientific contributions to the understanding of the effects and implications of government funded R&D activities. It will focus on recent theoretical, qualitative and quantitative studies and the exploration of different approaches to evaluate their effects. An applied workshop with policy makers will be organized on "What we know and what we should know about evaluation of technology policy" at the end of the conference.

Sessions:

- Methodology – Data and Evaluation Approaches
- Clusters, Networks & Spatial Dimension
- Evaluating the Rationales for RTD Policies
- Spillover Effects, Externalities and Productivity
- SMEs, Spin-Offs, NTBFs – Funding Gap in Innovation?
- The Evaluation of Infrastructure, and Institutions

Scientific Committee:

Spyridon Arvanitis (CH)
Robin Cowan (NL)
Susan E. Cozzens (USA)
Paul A. David (USA)*
Jonathan Eaton (USA)
Stefan Kuhlmann (D)
Georg Licht (D)
Albert N. Link (USA)
J. S. Metcalfe (GB)
Wolfgang Polt (A)
Bruno van Pottelsberghe(B)
David Storey (GB)
Nicholas S. Vonortas (USA)
* to be confirmed

Registration:
The number of participants is limited. The conference fee is €150,– (reduced €100,–) covering participation, luncheons and a conference reception. For participants who present a paper the conference fee will be covered by the organisers. Participants who do not intend to present a paper are requested to register not later than March 2003,30th. They will be accepted on a first come – first served basis.
For registration and further information concerning the conference and accommodation, please refer to the addresses given below. Papers and further information on the conference will be made available on the conference web page.

15th–16th May 2003

Tech Gate Vienna

Vienna, Austria

Conference Address:
TechGate Vienna, Donau-City-Straße 1, A-1220 Wien · (http://www.techgate.at)

Conference Organisers:

Plattform Forschungs– und Technologoieevaluierung GesbR & Joanneum Research
Klaus Zinöcker
Wiedner Hauptstr. 76
A–1040 Wien – Austria
E–mail: office@fteval.at
Website:www.fteval.at

Zentrum für Europäische Wirtschaftsforschung (ZEW)
Andreas Fier
P.O.Box 10 34 43
D-68034 Mannheim–Germany
E–mail:rd-eval@zew.de
Website:www.zew.de/rd-eval

bm:bwk
JOANNEUM RESEARCH
ZEW Zentrum für Europäische Wirtschaftsforschung GmbH Centre for European Economic Research

additionality has proved to be hard to define successfully, although much progress has been made in the last decade and we now have a good understanding of the possible ways in which net effects occur and could be measured. The ASIF report examines these issues in detail. Below is a brief review of some of the key points.

Net effects are a vital index of programme efficiency, effectiveness and appropriateness. Without knowledge of net effect, programme managers do not know whether the public money which they spend is generating social and economic benefits for society or is simply being wasted. Often however, additionality is considered alongside these other measures although it is in fact the basis from which all these other measures are derived.

Such a difference is often called the programme net effect, but the concept goes by a number of names, depending upon the context. The problems with understanding additionality begin with the question of whether one should look for additionality or net effect in terms of the extra work undertaken as a result of the subsidy – the extra commitment, or in terms of what is generated as a result of the action – the output. For example, if firms have undertaken a large amount of extra work above and beyond what they might have carried out without government support, then we might expect that the net effect is large, although this is not usually the case. This formulation is generally termed input additionality, or incrementality in some contexts, notably Canadian evaluation.

By contrast with the focus on how much difference has arisen at the level of inputs, one could also focus on the outputs of an action that has been supported and ask how much of the output is due to subsidy. If a firm, or a network of firms in collaboration, undertake work that has significant benefits when compared with the costs, then there is a significant net effect. But if a firm or a network undertakes a large amount of work but without significant benefit, then, while there appears to be a large amount of input additionality there is very little return, or what is called output additionality.

When programme intervention is assessed, it is often worthwhile to consider both input and output additionality. However, there is not a way of combining both measures in a composite measure of overall programme net effect, as the two concepts are not commensurable. The argument can be made that neither programmes with low input and low output additionality (no net effects) nor programmes with high input but low output additionality are worth running and that it only makes sense for programmes with low input additionality but high output additionality to be supported. In the figure below, these suggestions are outlined.

Figure 1. Input and Output Additionality – When to Run Programmes

	Input Additionality	
Output Additionality	High Input	Low Input
High	✓	✓
Low	X	X

Run=✓ Do not Run=X

One might therefore ask why we need a concept of input additionality at all if the concept merely identifies the firm's extra commitment to the project, given a certain level of government subsidy. The answer to this is that input additionality gives some indication about the help which firms receive which substitutes for or displaces the funds which the firm would have assigned to the project without support. And this substitution has further importance because it is a measure of the leverage effect of public money on the private resources of the firm and leverage of private money is an increasingly significant issue for research funding in the EU.

Of course, the calculation of net impact is very difficult: in fact it is a theoretical impossibility as the implementation of a programme removes the chance of ever making a comparison. However, the use of detailed case studies and mathematical techniques as exemplified in the evaluation of the United States Advanced Technology Programme provides means of making comparisons, give some indications as to the real difference between what actually happens and what would have happened in the absence of government action.

The following box indicates the four main types of additionality, which are discussed in depth in the ASIF Report and it shows the growing awareness of different dimensions to the net effect of RTD activities over the last decade.

Box 1. Different Types of Additionality

- Input additionality: whether the public action adds to, or substitutes for the agents inputs (usually financial);
- Output additionality: the counterfactual of whether the same outputs would have been obtained without policy action;
- Behavioural additionality: the differences to the agent's behaviour following the policy action, or its persistence beyond the action; and
- Cognitive capacity additionality: whether the policy action changes the different dimensions of the cognitive capacity of the agent.

Problems of Evaluation – some suggested methods

The ASIF Study reviewed a large number of evaluation methods and provides a number of recommendations on which evaluation techniques are appropriate for specific types of evaluation contexts. The review covers in detail the development of socio-economic impact assessment methodologies and derives a number of useful perspectives on their development. The report also includes a number of case studies of impact assessment in order to show the relevance of particular methods in particular contexts.

The first case study covers the use of a number of economic modelling techniques to understand the effects of RTD within the firm. Classical cost benefit analysis, option based methods and the detailed BETA Case Study Method are all applied to a single example to show how different methods can be used, their range, scope, feasibility and the extent to which such methods can be used together – i.e. their complementarity.

A second case study looks at the evaluation of the impacts of RTD within the context of service sector innovation. This example of a health care innovation demonstrates the importance of innovation champions in the development of particular technologies and the crucial role of standards and the existence of unintended impacts as technological developments are transferred between sectors.

A third case study looks at existing framework programme evaluation methods and principles and considers the impacts that have arisen. The report examines the use of ex post questionnaires, the role of external project evaluators and the combined use of questionnaires and mini case studies to focus on projects that have achieved significant socio-economic returns.

A fourth case study looks at the evaluation problems when assessing the impacts of the funding of social science research. The funding of this area has grown in recent framework programmes, but the attempt to define a causal link between publicly funded research and socio-economic benefit is a difficult problem to address. Preferred methods discussed in this section include the attempts to assess the solutions to research questions, attempts to measure changes in the level of interaction between actors identified as relevant to the solution of the research question, whether new research questions emerge, the creation of research resources and databases, the generation of new standards, protocols and methodologies and the performance of dissemination activities.

A fifth case study examines the role of framework programmes in the Accession and Cohesion countries and the evaluation of impact in such contexts. The case study reviews the impacts of framework programme involvement in such areas, finding that the changes to infrastructure and research and innovation cultures to be of major and lasting importance.

However, one of the main messages from the report is that there is no single method which can be used at all times and in all places to assess socio-economic benefits of RTD. Just as there are a number of views of additionality, so there are number of different types of methods for calculating the effects of programmes.

Evaluation in the European Research Area

The final section of the report examines the implications for evaluation in the European Research Area. In early 2000, the European Commission announced plans to transform the operation of the Framework Programmes. In the Commission's communication of early 2000, [Com (2000) 6 "Towards a European Research Area"] the new arrangements for the operation future FPs were outlined [specifically FP6]. In addition to the existing forms of support of specific targeted research programmes, coordination actions and specific support actions, the Commission proposed a further three new measures, comprising networks

of excellence, integrated projects, and application of Article 169 of the Treaty to national research programmes.

The justification of the new mechanisms was that European research funded through the Framework Programme should take place at a greater scale to give larger critical mass to research activities, greater coherence to European research activities, and greater scope for integration within Europe of different forms of research speciality. Whereas different national research systems had in the past worked separately with the risk, at times, of duplication of research activities, the European Research Area's new instruments are likely to bring about a new era of increasing collaboration between the research capabilities of member countries, with the Commission taking a leading role, coordinating and co-funding in a European nexus of research actors. However, there is no guarantee that the political willingness on the part of the many levels of stakeholders, particularly at national level, will allow the ERA to work as envisaged. In this section, the various evaluation challenges which arise from the likely implications of the introduction of the ERA concept are considered.

The first major challenges facing evaluators stem from the novelty of the new instruments. Firstly, there is the challenge of ensuring that the evaluation standards and practices that are used are appropriate methods with which to assess and measure the outcomes and impacts generated by this new range of modalities and instruments. This may be difficult as the scale of implementation of new instruments is likely to be conservative because of the risks involved, and this will provide limited scope for evaluation activities. Nevertheless, the requirement that the European Research Area is a success will ensure that evaluation is at the heart of discussion when the new instruments are introduced. Further challenges will come from those wishing to know how well the new instruments are operating: in addition to the relatively low numbers of projects operating under the new instruments, the new instruments are intended to have long term structuring effects. These will not be visible for some while, so the overall progress of the European Research Area in generating benefits and restructuring the research system will be hard to identify, at least in the short term. Furthermore, the success of the European Research Area is tied to closer integration between the research and higher education systems of the European Union – what is termed the European Higher Education Area – and developments such as this are likely to have more than a minor influence over the success of the ERA.

In terms of the specific instruments outlined, the proposal to establish networks of excellence takes an existing concept of operation known to work well at different scales and applies it at the level of the European Union. Particular challenges here will be the measurement of how well excellence is achieved; how the instrument contributes to national capabilities as well as to European ones; at what cost these achievements are made; and also how well the dissemination activities, which might lead to socio-economic impacts ultimately, are carried out. Comparisons with outside the EU will be possible here and the use of bibliometrics, including citation analysis, and network analysis techniques are likely to be favoured approaches to the identification of the outputs and status of these networks. Difficulties in the evaluation of impacts may arise when different kinds of support are offered to a network of excellence: when support for staff is given alongside infrastructural support, it may be difficult to make correct attribution of impact to effect. However, in networks of excellence, the emphasis is upon the scientific and academic case and a significant ex ante peer review process, with contributions from outside the EU, may be indicated to safeguard scientific quality and justify the creation of new centres.

Integrated projects present more difficulties to the evaluator. Here there is a significant need to demonstrate the value not only of the socio-economic impacts of the programmes themselves, including their additional impacts, but also that the structures themselves are effective and appropriate mechanisms by which European competitiveness is enhanced. A key concern is that such large systems achieve the integration of research and commercial organisations with lower levels of capabilities – principally SMEs, which currently under-perform, in research terms, their peers in the US.

There is also likely to be call for evaluation and monitoring at the portfolio level of the different integrated projects as such large entities effectively consolidate a large number of risks and if such projects fail, there is a significant loss of resources. A range of techniques to monitor the status of such projects in terms of their main effects and in terms of their progress towards project goals will be necessary. In any large system, there are likely to be parts that work well and parts that do not work so well. The larger the system, the more likely it is that there will be a bigger spread in the performance and effectiveness of the different elements that make up the whole. This makes the passing of a single one-dimensional judgment on the performance of large projects or projects problematic, as good networks may contain poor elements, and the less successful may contain many positive elements. Assessing larger structures at many levels appears the only safe way of knowing what such projects are achieving and giving insight into what can be done to ensure success in the future.

This discussion of monitoring and the reduction of risk raises the issue of what role ex ante and ex post evaluation are likely to play. It appears advisable that where new research instruments are to be launched, as they are in the case of the ERA, ex ante evaluation of project proposals should be emphasised to ensure as best a possible fit between projects and the allowed means (instruments) by which these projects will reach their goals. Ex ante evaluation for the new instruments is likely to have a greater priority, while for existing instruments, the emphasis is likely to remain on a range of ex post methods.

The introduction of the European Research Area is also likely to see demands that evaluation methodology address the issue of comparability of measurement and assessment processes between different areas, as national governments seek to prevent free-riding in terms of inputs into programmes. But there will also be demands from national governments that at the level of framework programme benefits, European-wide standards are introduced for reasons of fairness and equal returns (just retour) and so that the evaluation process itself can be trusted. These demands for transparency and comparability are likely to see evaluation systems adopt common standards and practices. There could also be pressure to adopt a larger number of quantitative measures, such as net prevent values and social rates of return which are used widely in the United States so as to provide highly visible comparisons for national legislatures.

Within the European Research Area, as policy instruments lead to the creation of a smaller number of larger projects, statistical methods which have been widely used in the assessment of input additionality become difficult to operate. In the US system of innovation funding, information about the performance of the ATP in terms of input additionality has resulted from detailed statistical analyses of applicants for programme funding. As the number of programmes falls in the European system, the use of such techniques becomes impractical, although econometric techniques can still be applied.

Because the ERA aims to change permanently the way in which research is carried out in Europe, with a corresponding change to the distribution of research activities and research expertise with individual country capabilities becoming less important than Europe-wide capabilities, a major challenge for those carrying out evaluation will be to ensure that the programme instruments for research, for innovation and for technological development are the most appropriate to the task; and in a wider sense, that the structures of the research systems can meet the many and diverse challenges which face Europe.

References

Buisseret, T. J., Cameron, H. M. and Georghiou, L. (1995): `What Difference does it Make? Additionality in the Public Support of R&D in Large Firms', International Journal of Technology Management, 10, pp. 587-600.

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

Author: Dr John Rigby
Research Associate
PREST
Harold Hankins Building
University of Manchester
Oxford Road
Manchester M13 9PL
tel.: + 44 (0) 161 275 5928
fax: + 44 (0) 161 273 1123
internal 55928
mobile 07811 852 826
john.rigby@man.ac.uk
http://les.man.ac.uk/PREST/

Jaime Rojo

Evaluation of RTD Policy Foundations: The Socio-economic Dimension

1. EVALUATION IN POLICY MAKING

All modern economies adopt more or less interventionist policies to support their science and technology system with the ultimate purpose of improving economic performance and social welfare. The justification for policy intervention in Research and Technological Development (RTD) results from a market failure in the private provision of RTD, due to the existence of externalities, market imperfections or asymmetric information. Under this situation, the social rate of return to RTD exceeds the private rate of return which leads from the social perspective to an underinvestment in RTD.

Evaluation plays a relevant role in the design and improvement of RTD policy by bringing accountability, transparency and rationalization into the policy making process. It does so by raising stake holders' awareness on policy impacts and improving the understanding on the rationales, implications and effects of policy intervention. Additionally, evaluation enlightens the policy formulation process by expanding the capacity to address new relevant policy topics and to understand policy limitations (Luukkonen 1997). These factors have contributed positively to the introduction of standardised evaluation requirements into programme management. The impact of the evaluation on policy definition will largely depend on the willingness of the evaluation sponsor to accept recommendations and critiques. Evaluation results are used for various purposes which include policy legitimisation, strengthening of RTD institutions and supporting sound policy reforms. The impact of the results of an evaluation should be judged on the basis of the policy debate generated by its recommendations and the capacity to highlight key issues at stake (Aubert 1998).

In the US, there is a larger tradition for conducting quantitative econometric evaluations. The policy recommendations stemming from these evaluations tend to be disruptive. In Europe, evaluation is used a means to review outcomes against stated objectives. It is driven by the need to comply with administrative regulation. Its recommendations generally focus on incremental changes in policy definition. However, the recently introduced reforms in national research and innovation systems have generated a greater reliance on evaluation results for allocating resources between RTD actors (Georghiou 2001, OECD 2002).

2. THE EVOLVING FRAMEWORK OF RTD EVALUATION

2.1 New RTD Policy Instruments

Evaluation practices need to co-evolve with the new instruments and rationales applied in RTD policy. New approaches should be devised to identify and measure more diffused intangible issues such as attractiveness, networking capacities and excellence that will be key determinants in ensuring that the best output is achieved out of RTD investments. For example, a larger emphasis will be given to policies that favour the development of an indigenous and sustained capacity to attract researchers from the rest of the world. The identification of these policies will require connecting measures of performance and innovation with indicators on working conditions, wages and career prospects.

The increase of multidisciplinarity as the mode of organisation of RTD and the multiplicity of actors involved requires a re-examination of traditional evaluation approaches. Furthermore, current evaluation practices have focused much more on evaluating behavioural additionality, i.e. the changes in the RTD players conduct induced by the policy (Georgiou and Roessner 2000).

Systemic approaches

Successful innovation increasingly depends on the efficiency of the organisational arrangements of the innovation system. Accordingly, the emphasis in evaluation is shifting towards approaches that allow to assess the design and capacity of the innovation system to achieve its best socio-economic results. In the knowledge-driven economy, RTD is no longer an activity performed in isolation; factors such as networking, collaboration, multidiciplinarity, access to information and mobility all becoming crucial in the process. This requires evolving towards systemic evaluation approaches that allow to assess the capabilities of complex institutional set-ups. Evaluation of universities and research institutions should be conducted taking their context into consideration and encompassing the various objectives defined in their mission statements such as advancing knowledge and promoting education. In contrast, traditional project or programme evaluations have dealt with a limited set of objectives.

Human capital

The centrality of human capital in knowledge generation and diffusion implies that evaluation should pay much more attention to the mobility of skilled workers as well as to their motivations and incentives. Tacit knowledge as embodied in human capital has become a key aspect for innovation. The strategic relevance of this factor has increased as a consequence of the globalisation and increased factor mobility trends. The mobility of researchers between academia and industry and their geographical mobility have become fundamental elements for enhancing the process of knowledge generation and diffusion. Evaluation can shed light on understanding the development and motivation leading individuals to follow research careers.

Research Excellence

Evaluation can provide indicators to support the mapping of research excellence. The expanded use of benchmarking approaches in evaluation reflects the recognition of the large effects that policy exerts on the behaviour of supported firms and individuals. Benchmarking allows to identify and compare practices to improve policy performance by adopting the perspective of the users of the evaluation results. However, the various indicators used in benchmarking should be connected to a formal theory, linking the policy intervention to achieved policy outcomes.

The trend towards larger size RTD projects requires a more rigorous use of evaluation approaches to analyse the degree of accomplishment of project objectives and achieved results. The involvement of a growing multiplicity of actors interacting at different stages of the RTD process highlights the relevance of monitoring and evaluating the project results against its stated objectives. In large projects, there is scope for a more rigorous application of quantitative techniques, such as cost-benefit analysis and econometric methods that will provide an estimate of the orders of magnitude involved. As a larger dimension of projects is likely to mobilise a larger amount of resources, the impact achieved will be easier to trace in the economy. The decision of financing large research facilities is likely to rely increasingly on supporting economic indicators and criteria, although, quite likely, the main element to guide investment will continue to depend on scientific excellence criteria. The significant increase in the costs of research implies that the competition for funds among disciplines will require a more regular and standardised use of methods of economic assessment to evaluate the potential benefits and costs of alternative investment opportunities.

2.2 Expanded Rationale of Public Intervention

Recent policy instruments are implicitly or explicitly founded upon an expanded rationale for government intervention in which the market does not build the 'bridges' necessary to link actors within the innovation system, creating barriers to innovation, even though such bridges would generate profits. In this case the role of the state is that of providing bridging institutions linking the research and commercialisation phases of innovation, stimulating technology transfer and the take up of new technologies by firms. The new forms of institutional set-ups include incubators, technology parks, venture capital, technology transfer offices, research joint ventures and research spin-offs (Branscomb, Kodama and Florida 1999, OECD 2002).

The closer interaction between private and the public sector require devising better methods to assess collaboration linkages, networking capabilities and new forms of RTD policy. Network and cluster analysis provide insights on the connections between research actors. Novel approaches have been used to benchmark private-public partnerships and industry-science relations.

Evaluation might help in clarifying the impacts of recent policy regime changes in the stimulation, diffusion and transfer of knowledge and S&T results to the marketplace.

2.3 Societal Dimension of RTD Policy

Nowadays, as any other social institution science is viewed as requiring both accountability and the need to demonstrate its expected returns (Nowotny 1997). Evaluation practices have to deal with the sometimes conflicting rationales applied in policy making and incorporate the demands and expectations of a wider group of interested individuals and communities.

After a recent upsurge of public mistrust in science, evaluation by providing evidence and demonstrating the positive effects of science to society can help to restore public trust in science and technology issues.

Evaluation methodologies have to evolve to incorporate the increasing relevance given to socially agreed norms and ethical values in the context of scientific and technological matters (e.g. biodiversity, environment, sustainable development, genetically modified organisms). This will require a strengthened coordination and consistency of the views shared by decision makers.

2.4 Structuring Effects of RTD Policy

A relevant aspect in evaluation refers to the structuring capacity of RTD policy in promoting regional development and economic cohesion. As with instruments for innovation, there is a need for rigorous studies on best practices of technology transfer schemes, identifying good practice examples of knowledge transfer at regional and local levels. In this context, it is relevant to evaluate the success of RTD policies aimed at strengthening the catch-up of lagging regions. The clarification of policy rationales and impacts could lead to a more coherent implementation of regional, national and supranational research activities. This requires better information on the relative success of the different policy instruments and a better understanding of their operation.

2.5 Selected Experiences on RTD Policy Evaluation

The US has accumulated a wide experience in RTD evaluation, ranging from the evaluation of specific programmes such as the Advanced Technology Programme (ATP), Cooperative Research and Developments Agreements (CRADAS), extension programmes, Small Business Innovation Research (SBIR) Programme to specific regulations such as the Bay-Dole Act and the Government Performance Results Act (GPRA) and concrete actions such as the creation of technology transfer offices and liaison offices.

In Europe, a wide knowledge has developed among countries with different evaluation traditions (Callon, Larédo and Mustar 1997, Kuhlmann and Shapira 2001). For example, in the UK the funding of RTD programme by government requires that government agencies follow the Rationale, Objectives, Appraisal, Monitoring, Evaluation and Feedback (ROAMEF) approach. Moreover, reforms and financial pressures have pushed universities and public laboratories to enter the market for contract research and evaluation of research departments is now a common element of the UK system (Georghiou 2001). In Germany, the major research and technology organisations including institutions such as the Max Planck, Fraunhofer Gesellschaft and Helmholtz have gone through system evaluations. France has developed a strong tradition in the evaluation of institutions and research capacities. The scope of evaluation has expanded to cover other institutional set-ups such as Research Councils. The OECD has sponsored national programmes reviews of its member countries.

At a European wide level, the European Commission has accumulated a consolidated experience in the evaluation of the Framework Programme and its specific RTD programmes (Airaghi et al., 1999; Georghiou, Rigby and Cameron 2002). It has developed a quite unique, well organised and structured evaluation system including (Fayl et al.,1999): (i) project appraisal or ex-ante project and programme evaluation; (ii) external annual monitoring and mid-term evaluation of projects and programmes implementation (European Commission, 2001); (iii) horizontal or thematic evaluations – e.g. SME, Information Society, and (iv) 5-year assessment of whole programme, (European Commission 1997, European Commission 2000a), and each of the specific RTD programmes, including a communication of the of the Commission commenting on the conclusions of the assessment (European Commission 2000b) as well as a questionnaire survey on a selected sample of projects. Most EU Member States also conduct country evaluation of their own participation in the Framework Programme.

3. EVALUATION METHODOLOGIES

Evaluation methodologies rely on modelling frameworks that describe the process of research and technological innovation. The design of evaluation methodologies based on economic theoretical foundations allows to relate output measures with the broader impact of the policies on economic welfare. Many times, for simplifying purposes, evaluation approaches adopt the linear model of innovation framework. However, for some purposes and within certain contexts, more complex models of innovation are used in evaluation. For instance, the measurement and description of the creation of knowledge clusters, industry-science relationships and private-public collaborative agreements, adopt more complex modelling frameworks of national systems of innovation, chain linkage models of innovation and models of creative destruction (OECD 2002).

There is a growing literature that focuses on the methodologies employed in RTD evaluation. The methods used in research impact evaluation – e.g. bibliometrics, peer review – are explored in Bozeman and Melkers (1993). The specificities of policy evaluation are examined in Papaconstantinou and Polt (1997) while the methodologies used to evaluate the socio-economic impacts of RTD policies are emphasised in Polt and Rojo (2002). The variables of interest in policy evaluation, outcome and impact, are captured indirectly by analysing how the outcomes of intervention change as a result to changes in the relationships and amounts of inputs and outputs. As evaluation approaches recognize, institutional variables are a fundamental factor shaping the effects of policy interventions. The different available evaluation methodologies, specially the qualitative ones, indicate the relevance of context dependence in evaluation.

The use of quantitative methods combined with performance indicators in RTD policy evaluation allows to capture involved dynamics providing good estimates of output and impact of public intervention. Policy makers could then make use of these impact estimates as a means to legitimise their actions and as supporting evidence of the policy rationale. Econometric studies provide insights on the effects of RTD on economic productivity and the magnitude and source of spillovers. Cost-benefit analysis provides a useful synthesis of detailed cost and benefits of a programme. Microeconometric techniques allow to explore the leverage effects of the intervention, the crowding-out of private RTD and the private returns to RTD. Survey studies in combination with statistical techniques allow to study more qualitative aspects of firm innovation and concentrate on changes experienced by the variables. Qualitative evaluation methods such as interviews and case studies provide policy makers with more rich and detailed insights on the multiple effects of policy intervention, helping to improve and clarify the processes, instruments and behavior induced by science and technology policy.

Evaluation methodologies have been adapted to capture the benefits of policies not directly quantifiable in monetary terms – for example, in the areas of health, environmental sustainability, consumer protection. The techniques developed to measure the returns in these fields include contingent valuation and revealed preferences. Hedonic prices allow the adjustment of price series to reflect changes in the quality of products. This is particularly relevant issue in high technology products subject to rapid quality changes. Methodological approaches have been developed to cope with the uncertainty and risk inherent in RTD process. This has lead to the use of decision theory in a cost-benefit analysis context. Uncertainty considerations can be introduced in ex-ante evaluation to provide a more accurate assessment of the return on RTD investment.

4. LESSONS FOR RTD POLICY EVALUATION

The need of an enlarged interconnection between policy-making and evaluation at all levels of decision-making is emerging. However, evaluation approaches are still deeply rooted in independent national administrative traditions which show strong cultural and historical path dependence patterns. There is scope for sharing good practice experiences and for advancing in the standardisation and harmonisation of shared evaluation principles and approaches across different evaluation cultures and traditions.

The role of the government in steering the innovation process is becoming ever more subtle, reducing it scope to the provision of those framework conditions that stimulate the innovation process. Successful innovation is increasingly the result of a consensus building approach achieved in a multi-actors multi-level environment where the government adopts the restrained role of acting as a mediator in the process. In this context, the use of evaluation approaches might help to create a consensus between agents sharing divergent positions on policy matters.

The series of recommendations detailed below were extracted in the context of the thematic network on the Socio-Economic Evaluation of Public RTD policies (EPUB) for improving the evaluation of RTD policies (Farenkrog et al. 2002, Polt and Rojo 2002):

- Stimulate flexibility and combination of approaches in evaluation, devising methods to cope with the emergence of new policy instruments. Evaluation should be flexible enough to adapt to the challenges posed by the need to assess the socio-economic impacts of the new forms of RTD policy based in extended policy rationales. The evaluation of more complex institutional set-ups requires devising systemic evaluation approaches.
- Support the use of more analytically-based and data driven techniques in evaluation to bring transparency and credibility to the evaluation recommendations. Making explicit the assumptions made and providing a quantification of the effects, stimulates the discussion, replication and validation of evaluation results by third parties. Conducting evaluation studies at different levels of data aggregation permits to get a more complete picture of policy impacts and reinforce evaluation results.
- Improve the quality of the data used in evaluation studies. Best fitted econometric techniques will fail if used in combination with poor quality data. The process of data collection for programme management purposes should take into consideration the needs of evaluation. The process of data harmonisation and update is a particularly critical issue in evaluation as most evaluations require some sort of comparisons – e.g. before/after, participant/non-participant, etc.
- Clarify the objectives of the policy intervention internalising the sometimes diverging rationales as they will guide the evaluation process and will assist in the formulation of the type of questions the evaluation tries to answer. The evaluation operational approaches should be adapted to

the questions that the evaluation tries to answer, considering the objectives, data and indicators available. Evaluation as an operative discipline depends very much on practice and application.

- Respect some basic methodological guidelines when performing quantitative evaluations. For example, introduce the use of a control group to correctly estimate the difference the programme makes on the performance of the participant firm and investigate the characteristics of the rejected firms. Conduct detailed case studies to measure the returns of the reduced number of projects that lead to large payoffs. Adopt benchmarking practices to evaluate new forms of performing RTD such as private-public partnerships and industry-science relations. Use econometric approaches to investigate the sources of RTD spillovers and the geographical distribution of knowledge and localisation effects.
- Implement new approaches that allow to evaluate the outcome and impact in new scientific disciplines where the application of well established peer review methods would not help. This would probably require the use of extended expert panels and users' surveys involving industry, community groups and associations, technology users and suppliers that help to ensure the social acceptability of the new technology.
- Promote acceptability and learning by incorporating relevant stakeholders early in the process and diffusing evaluation results to the wider public. Strategy policy definition could be supported by connect evaluation results to other sources of RTD intelligence such as foresight, technology assessment and technology forecasting. Independence in evaluation could be promoted by not penalizing reporting of negative findings. Diversification of funding sources for evaluation studies could assist in this purpose.

5. SOME OPEN QUESTIONS

An examination of the literature on RTD policy evaluation allows to identify a series of open questions that will require further investigation if we are to improve our understanding of RTD policy and evaluation approaches.

- What are the policies influencing innovation? Little is still known about the effects that other policies, besides RTD policy, have in the efficiency of the innovation system. This includes devising methods to study the likely influences that trade, financial, labour market, competition, regulatory and structural policies have on innovation efficiency and institutions motivation.
- How to achieve a better appreciation of the aggregate picture on policy effects? There is a limited understanding of the effects generated by policy intervention at the macro level due to the difficulties of data aggregation and the limitation of methodologies to capture the indirect effects. This requires improving the capacity policy level data from project and programme data.
- How to maximize the returns from RTD? What explains the wide variation in returns of RTD across sectors? What are the enabling technologies providing the largest productivity increases in the different economic sectors? Is it better to fund RTD diversity or to concentrate research funding?
- How to improve the measurement of the RTD spillovers? Improve the understanding of the sources of RTD spillovers. The intra and inter industry indirect effects of a programme are not well captured with currently available methods. At the micro level, control group approaches have difficulties to capture the induced external effects the programme has on the conduct and performance on non-participants.
- Are RTD investments subject to diminishing returns? This is linked to the debate on the productivity paradox and the implications of the new growth theory. Part of the problem lies in our limited knowledge on how to capture quality improvements in new and high technology products.
- What are the various dimensions in which RTD impacts firms' output? The use of cross country longitudinal data might help to improve the understanding of the impact of RTD as a productivity enhancing mechanism and as a mechanism to improve the capacity of firms to learn from others knowledge.
- How to improve the assessment of social impacts? How to measure the quality of life without introducing subjectivity into the assessment? Results of these studies are quite sensitive to the value of human life and the discount rate applied. How aspects of consumer protection should be incorporated in defining strategic policy options? How should the precautionary principle be incorporated in technology risk assessment?
- How to improve measurement in RTD? How to capture the return of basic research? How to assess the economic impact of increasingly relevant scientific and technological areas – e.g. biotechnology?

6. CONCLUDING COMMENTS

The evaluation of the socio-economic impacts of science and technology policy is becoming an increasingly relevant topic in policy making. At present, most RTD policies define specific socio-economic goals against which achieved results should be evaluated. This is not an easy task given the uncertainty and intangible character of science and technology. At the same time, evaluation practices have to evolve in order to cope with the new extended rationales applied in science and technology policy.

Methodological and theoretical advances as well as improvements in data availability have contributed to improve the practice of evaluation. However, no unique best evaluation approach exists. Most evaluations will benefit from a combined use of various methods which provide complementary information on the various dimensions of policy effects.

Despite the availability of a number of evaluation methods, there is scope and need to look for further methodological improvements in evaluation. At present, consistent evaluations can be conducted at the project level, but undoubtedly more thorough evaluations at programme or policy level will require advances in knowledge both in the causal relations between inputs and outputs and in the way of arriving to meaningful procedures for measuring and aggregating these outputs.

References

Airaghi A., Busch, N., Georghiou, L., Kuhlmann, S., Ledoux, M., van Raan, A. and J. Viana Baptista (1999), Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes, ETAN, Commission of the European Communities, January 1999.

Aubert, J. (1997), "On the OECD Experience of Country Reviews" in G. Papaconstantinou and W. Polt (eds.) Policy Evaluation in Innovation and Technology: Towards Best Practices, Paris: OECD.

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

Branscomb, L., Kodama, F. and R. Florida (eds.) (1999), Industrialising Knowledge: University-Industry Linkages in Japan and the United States, Massachusetts: MIT Press.

Callon, M., Larédo, P. and P. Mustar (1997), The Strategic Analysis of Research and Technology- Evaluation of Programmes, Paris: Economica.

Fahrenkrog, G, Polt, W., Rojo J., Tubke, A. and K. Zinöcker (eds.) (2002), Research and Technological Development Evaluation Toolbox, IPTS Technical Report, EUR 20382 EN, European Commission September

European Commission (1997), Five-Year Assessment of the European Community RTD Framework Programs, 1991-1995, EUR 17644, Brussels: European Commission.

European Commission (2000a), Five-Year Assessment of the European Union Research and Technological Development Programs, 1995-1999, EUR 19426, Brussels: European Commission.

European Commission (2000b), Comments of the Commission on the Conclusions of the RTD Framework Programmes 5-Year Assessment, COM(2000)659, Brussels: European Commission.

European Commission (2001), 2000 Annual Monitoring Report on the RTD Activities Conducted under the EC and Euratom Framework Programmes, Brussels: European Commission.

Fayl, G., Dumont, Y., Durieux, L., Karatzas, I. and L. O'Sullivan (1998), "Evaluation of Research and Technological Development Programmes: A Tool for Policy Design", Research Evaluation,7(2), 93-97.

Georghiou, L. (2001), "The United Kingdom National System of Research" in P. Larédo and P. Mustar (eds.) Research and Innovation Policies in the New Global Economy, Cheltenham: Edward Elgar. 253-296.

Georghiou, L., Rigby, J. and H. Cameron (eds.) (2002), Assessing the Socio-Economic Impacts of the Framework Programme, PREST, Manchester.

Georghiou, L. and D. Roessner (2000), "Evaluating Technology Programmes: Tools and Methods", Research Policy, 29, 657-678.

Kuhlmann, S. and P. Shapira (2001), Learning from Science and Technology Policy Evaluation- Proceedings from 2000 U.S.- European Workshop, Bad Herrenalb, Germany.

Luukkonen, T. (1997) "The Increasing Professionalisation of the Evaluation of Mission Oriented Research in Finland: Implications for the Evaluation Process" in G. Papaconstantinou and W. Polt (eds.) Policy Evaluation in Innovation and Technology: Towards Best Practices, Paris: OECD.

Nowotny, H. (1997), "New Societal Demands" in R. Barré,

M. Gibbons, J. Maddox, B. Martin and P. Papon (eds.), Science in Tomorrow's Europe, Paris: Economica.

OECD (2002), Benchmarking Industry-Science Relationships, Paris.

Papaconstantinou, G. and Polt, W. (eds.) (1998), Policy Evaluation in Innovation and Technology: Towards Best Practices, Paris: OECD.

Polt, W. and Rojo, J. (Editors), Handbook on the Evaluation of Research and Technology Policies: Concepts, Tools and Indicators. Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing (forthcoming).

Author: Dr. Jaime Rojo
Universidad Politécnica de Valencia
Instituto de Gestión de la Innovación y del Conocimiento - INGENIO
Edificio 14, Camino de Vera s/n
E- 46071 Valencia
tel.: +34 96387704
fax: +34 963877991
jrojo@ingenio.upv.es
http://www.ingenio.upv.es/inicio.html

PLATTFORM ft eval
Forschungs- und Technologieevaluierung

Spyros Arvanitis

Buchbesprechung / Book Review: Staatliche Förderung industrieller Forschung in Deutschland (Andreas Fier)

The book to be discussed here has a clear structure: right from the beginning focuses to two general hypotheses which build the background for the two main parts of the study. Hypothesis I is: the more extensive the public financial R&D support of firms is, the stronger is also the impact on enterprise performance. (I would add, given a constant efficiency of public funds). Given the tendency of decreasing public support budgets the idea behind this hypothesis is an elementary one: also in R&D there cannot be any output without (financial) input. Hypothesis II reads: the impact of public support is complementary to firms' own efforts in R&D. We may argue that this complementarity has always been the intension of policy makers, but the real issue is, at what extent has this goal ever been attained in reality? This is an exciting question which attracts at the moment the attention of many economists as well as policy makers interested in an efficient use of public funds in times of (once more) growing budget deficits in most European countries.

Part 1 deals with the evolution of public R&D support in Germany since the end of the Second World War and results to a comprehensive history of German technology policy based on a sound empirical ground; this is a more history-oriented, descriptive and qualitative piece of work based on a great amount of empirical information. The author distinguishes several stages of the evolution of the German public R&D support policy and develops a qualitative stage model which helps to understand continuance as well as changes in support policy, discusses the evolution of goals, instruments and processes of R&D policy in political and economic context, compares the growth of public expenditures as a percentage of GDP for a number of countries and analyzes a large database on supported projects, expenditure volumes, focal points of R&D support and change of research priorities for a period of about twenty years (1980 -1998) (chapters 2, 3, 4 and 5). From the very interesting material described and presented in this part of the study we can pick out here only some highlights. The stage model in chapter 2 seems quite plausible to the interested non-specialist on modern German economic history: six clearly distiguishable periods are postulated beginning with the "rebuilding" ("Wiederaufbau") of the economy in the fifties and ending with the accelerated "globalization" of the economy in the nineties. An international comparison of public expenditure in R&D as a percentage of GDP in chapter 4 shows that Germany retained a top position with respect to non-military R&D expenditures between 1989 and 1998 spending between 0.77% (1998) and 0.93% (1989) of its GDP for this purpose (ahead of the USA, France, Japan and the United Kingdom). Chapter 5 offers, among many other interesting things, an impressive overview of the shift of public support of energy technologies at the beginning of the eighties (56.5% of the funds in 1981/82; 18.8% in 1997/98) to research in information amd communication technologies, biotechnology and material technology at the end of the nineties (11.6% of the funds in 1981/82; 42.4% in 1997/98).

Part 2 of the book deals with the problem of a sensible quantitative estimation of the economic impact of R&D public support (chapters 6, 7 and 8). This is the more analytical and quantitative part of the study. It is worth looking more carefully at the methodology and the results of this study because it is one of the first German studies using this kind of analysis, which became known to a wider audience of economists and policy makers at the latest since James Heckman received the Nobel Prize two years ago for the development and application of a new econometric methodology for the empirical investigation of the economic impact of policy interventions.

The policy instrument to be evaluated was the direct R&D-oriented promotion projects ("direkte Projektförderung") of the German Federal Ministry of Education and Research (BMBF). A database which contained data for 3136 firms, from which 297 were publicly supported in the period 1992 -98, was constructed by matching firm data from the ZEW enterprise panel with financial R&D support data of the Federal Ministry. The policy goal, whose extent of attainment had to be econometrically investigated, was the enhancement of the overall R&D intensity of the supported firms through the public support. In order to check the robustmess of analysis results two methods were applied: an econometric estimation of a R&D-intensity equation (which did not include controls for selectivity biases with respect to the supported firms) and a comparison of pairs of structurally similar supported and non-supported firms based on the "propensity score" matching method (which ex definitionem control for selectivity biases with respect to the supported firms). It is a considerable merit of this investigation that an important precondition, particularly for the appropriate application of the "matching" method, was fulfilled in this study: the availability of a series of variables which could serve as comparison criteria (as well as exogen

variables for the R&D intensity model). The author was able to collect data for about a dozen relevant variables such as number of employees, firm age, market share, capital intensity, degree of product diversification, export and import share, etc. The results showed that: a) there are substantial positive effects of public support projects on the R&D intensity of supported firms; the estimated gross effect was 2.8 to 4.4 percentage points, meaning that an increase of one percentage point of support would cause an increase of the R&D intensity by 2.8 to 4.4 percentage points; b) this effect is at largest for SME, "medium-sized" financial support and for research in material technology as well as research in physico-chemical technologies.

In sum, this is a well-conceptualized and well-structured study which I warmly recommend for reading for everyone interested in the economic effects of technology policy, not only in Germany.

This book is the doctoral dissertation of the author at the University of Munich. Andreas Fier is now a Senior Researcher in Zentrum für Europäische Wirtschaftsforschung in Mannheim, Germany.

Author: Spyros Arvanitis
Swiss Federal Institute of Technology Zurich
ETH Zentrum
CH-8092 Zurich
Phone +411/632'51'68
E-Mail arvanitis@kof.gess.ethz.ch
Website www.kof.ethz.ch

Andreas Fier: Staatliche Förderung industrieller Forschung in Deutschland. Eine empirische Wirkungsanalyse der direkten Projektförderung des Bundes, ZEW Wirtschaftsanalysen, Band 62, 303 Seiten, Nomos Verlagsgesellschaft, Baden-Baden, 2002. ISBN 3-7890-8087-X

Der Newsletter der Plattform Forschungs- und Technologieevaluierung GesbR
ist ein unregelmäßig erscheinendes offenes Forum zur Diskussion methodischer
und inhaltlicher Evaluierungsfragen in der Forschungs- und Technologiepolitik.
 ISSN: 1726-6629

Herausgabe und Versand:
Dr. Rupert Pichler, Dr. Michael Stampfer, Mag. Klaus Zinöcker
Plattform Forschungs- und Technologieevaluierung GesbR,
A-1040 Wien, Wiedner Hauptstraße 76. office@fteval.at

Für den Inhalt dieser Ausgabe verantwortlich:
Plattform Forschungs- und Technologieevaluierung GesbR
Mag. Klaus Zinöcker

PLATTFORM Forschungs- und Technologieevaluierung fteval

PLATTFORM FORSCHUNGS- UND TECHNOLOGIEEVALUIERUNG

Die Plattform Forschungs- und Technologieevaluierung GesbR ist eine Initiative des österreichischen Bundesministeriums für Verkehr, Innovation und Technologie, des Bundesministeriums für Bildung, Wissenschaft und Kultur, des Bundesministeriums für Wirtschaft und Arbeit, Joanneum Research, WIFO, ARC Seibersdorf research, der KMU Forschung Austria, der Technologie Impulse Gesellschaft, Technopolis der WWTF sowie den Fonds FFF und FWF. Im Rahmen der Plattform werden Themenstellungen zur Forschungs- und Technologieevaluierung erarbeitet und – z.T. unter Einbeziehung namhafter ExpertInnen – in einem Fachkreis diskutiert. Der Newsletter beinhaltet Fachbeiträge zu Fragen der forschungs- und technologiepolitischen Evaluierung. Die Herausgabe erfolgt in zeitlicher als auch inhaltlicher Abstimmung mit Plattform-Veranstaltungen, um die Synergiewirkungen eines breiten Austauschforums zu nutzen.

Plattform Forschungs- und Technologieevaluierung GesbR.
Wiedner Hauptstr. 76 A-1040 Wien, Austria
www.fteval.at