JOANNEUM RESEARCH

INSTITUTE OF TECHNOLOGY AND REGIONAL POLICY
INTEREG **RESEARCH REPORT SERIES**

ISSN 1810-6307

InTeReg Research Report No. 02-2002

RTD EVALUATION TOOLBOX

ABRIGED VERSION WITH SELECTED EXCERPTS

Wolfgang Polt, Klaus Zinnöcker

June 2002

JOANNEUM RESEARCH Forschungsgesellschaft mbH – Institut of Technology and Regional Policy (InTeReg)

Vienna Office:
Wiedner Hauptstraße 76
A-1040 Vienna, Austria
Phone: +43-1-581 75 20
E-Mail: intereg@joanneum.at

Graz Office:
Elisabethstraße 20
A-8010 Graz, Austria
Phone: +43-316-876 1488
E-Mail: intereg@joanneum.at

epub

RTD Evaluation Toolbox

[ABRIDGED VERSION WITH SELECTED EXCERPTS]

JUNE 2002

ipts

JOANNEUM RESEARCH

This publication is part of a project called "Socio-Economic Evaluation of Public RTD policies (EPUB)" which has been awarded financial support by the European Commission through contract nr. HPV1-CT-1999-00005 under the 5th Framework Programme of the European Community for Research, Technological Development and Demonstration Activities (1998 to 2002), and its specific programme "Improving the Human Research Potential and the Socio-Economic Knowledge Base" ("Strategic Analysis of Specific Political Issues"). The authors are solely responsible for this publication, which does not represent the opinion of the Commission. Neither the Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of the information contained in this publication."

Socio - Economic Evaluation of Public RTD Policies (EPUB)

The assessment and development of methods and mechanisms to evaluate the socio-economic impact of public Research and Technological Development (RTD) policies constitutes a highly relevant component for the efficient articulation and improvement of science and technology policy.

The EPUB thematic network constitutes a pilot action aiming at supporting European RTD policy making and improving efficiency in RTD policy actions. The network main objective of is to provide a better understanding of the dynamics of science and technology by helping to investigate the socio-economic impact of private/public RTD policies at a micro/macro level. In this context, a special emphasis is given to analyse and review the rationales, concepts, methodologies and indicators generally applied in RTD evaluation and to investigate the implications for evaluation of new S&T policy approaches. The network also facilitates the sharing and diffusion of knowledge, techniques and tools applied in RTD evaluation.

The RTD Evaluation Toolbox constitutes a compendium of current methods and practices used in the field of RTD policy evaluation. It is hoped that this effort evidences the close interaction that exists nowadays between evaluation and policy making and the possibilities it brings for experiencing mutual learning and cross fertilisation.

Contact

European Commission
Joint Research Centre - Institute for Prospective Technological Studies
Edificio Expo W.T.C., Calle Inca Garcilaso s/n
E-41092 Sevilla, Spain

Gustavo Fahrenkrog
Tel.: + 34.95.448.84.50
Fax: + 34.95.448.83.59
e-mail: gustavo.fahrenkrog@jrc.es

Alexander Tübke
Tel.: + 34.95.448.83.06
Fax: + 34.95.448.83.26
e-mail: alexander.tubke@jrc.es

Joanneum Research
Institute of Technology and Regional Policy
Institut für Technologie- u. Regionalpolitik
Wiedner Hauptstraße 76
A-1040 Vienna, Austria

Wolfgang Polt
Tel.: + 43.1.581.75.20.2811
Fax: + 43.1.581.75.20.2820
e-mail: wolfgang.polt@joanneum.at

Jaime Rojo
Tel.: + 43.1.581.75.20.2815
Fax: + 43.1.581.75.20.2820
e-mail: jaime.rojo@joanneum.at

Klaus Zinöcker
Tel.: + 43.1.581.75.20.2827
Fax: + 43.1.581.75.20.2820
e-mail: klaus.zinöcker@joanneum.at

Network Members

	European Commission, Joint Research Centre DG Institute for Prospective Technological Studies Gustavo Fahrenkrog / Alexander Tübke Edificio World Trade Center, Isla de la Cartuja s/n 41092 Sevilla, Spain Phone/fax: +34 95 448 83 61 / +34 95 448 83 59 e-mail: gustavo.fahrenkrog@jrc.es /alexander.tubke@jrc.es
	Joanneum Research Forschungsgesellschaft mbH Wolfgang Polt/ Jaime Rojo / Klaus Zinöcker // Markus Michaelis Wiedner Hauptstrasse 76, 1040 Wien // Steyrergasse 17, 8010 Graz/ Austria Phone/ fax: +43 1 58 17 520 2811/ +43 1 58 17 520 2820 e-mail: wolfgang.polt@joanneum.at / jaime.rojo@joanneum.at/ klaus.zinoecker@joanneum.at / markus.michaelis@joanneum.at
	The CIRCA Group Europe Limited Eamonn Kinsella Upper Pembroke Street 26, 2 Dublin, Ireland Phone/fax: +353 1 662 5895 / +353 1 637 3986 e-mail: eamonn.kinsella@circa.ie
	University of Sussex SPRU (Social Policy Research Unit) Pari Patel Sussex House, Falmer, BN1 9 RH, Brighton, United Kingdom Phone/fax: +44 1273 68 67 58 / +44 1273 68 58 65 e-mail: parip@sussex.ac.uk
	The Tavistock Institute of Human Relations Elliot Stern 30 Tabernacle Street, EC2A 4UE, London, United Kingdom Phone/fax: +44 171 417 0407 / +44 171 417 0567 e-mail: e.stern@tavinstitute.org
	The Victoria University of Manchester PREST (Policy Research in Engineering Science and Technology) Luke Georghiou Oxford Road, M13 PL Manchester, United Kingdom Phone/fax: +44 161 2755921 / +44 161 2731123 e-mail: luke.georghiou@man.ac.uk
	ZEW (Zentrum europäische Wirtschaftsforschung) (Centre for European Economic Research) Georg Licht 103443 L7,1 68034 Mannheim, Germany Phone/fax:+ 49 621 1235177/ +49 621 1235170 e-mail: licht@zew.de

ULB	Université Libre de Bruxelles Unité d'économie spatiale et de la technologie Dulbea – CP 140 Henri Capron Avenue F. Roosevelt 50, 1050 Brussels, Belgium Phone/fax: +32-2 6504141 /+32-2 650 38 25 e-mail: hcapron@ulb.ac.be
	Boston University and NBER (National Bureau of Economic Research) Jonathan Eaton 63 Westcliff road, 02493 Weston, MA, United States Phone/fax: +1 617 353 4142 /+ 1 617 353 4999 e-mail: jeaton@bu.edu
ISI Fraunhofer Institute Systems and Innovation Research	Fraunhofer Gesellschaft zur Förderung der angewandten Forschung Fraunhofer Institute Systems and Innovation Research (FhG-ISI) Stefan Kuhlmann Breslauer Strasse 48, 76139 Karlsruhe, Germany Phone/fax: +49 721 6809 170/4 / +49 721 6809 260 e-mail: sk@isi.fhg.de
ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich	ETH (Eidgenössische Technische Hochschule Zürich) Swiss Institute for business cycle research industrial economics research group Heinz Hollenstein Weinbergstrasse 35, ETH Zentrum, CH-8092, Zurich, Switzerland Phone/fax: +41 1 632 59 29 / +41 1 632 12 18 e-mail: hollenstein@kof.reok.ethz.ch
Consiglio Nazionale delle Ricerche	Consiglio Nazionale delle Ricerche Istituto di Studi sulla Ricerca e Documentazione Scientifica Giorgio Sirilli Via Cesare de Lollis 12, 00185 Roma, Italy Phone/fax: +39 0644879214 / +39 064 463836 e-mail: sirilli@www.isrds.rm.cnr.it
MERIT	University of Maastricht MERIT (Maastricht Economic Research Institute on Innovation and Technology) Luc Soete PO Box 616, Tongerstraat 49, 6200 MD, Maastricht, Netherlands Phone/fax: +31 43 3883891 / +31 43 3216518 e-mail: luc.soete@algec.unimaas.nl

Contents

List of Tables

List of Figures

Executive Summary

All modern economies adopt policies to support their science and technology system aiming to improve economic performance and social welfare. This increasing relevance of science and technology in shaping the forces driving our societies and economies is forcefully reflected in the Presidency's conclusions of the European Council held in Lisbon on March 2000 where the Heads of Government and State agreed on the new strategic goal of the EU to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable growth with more and better jobs and greater social cohesion. Fulfilling this challenging strategy requires a clear commitment on the part of governments to discern the socio-economic impacts of science and technology policies.

Evaluation is a key decision support tool providing policy makers with a better understanding of policy results, allowing learning from past experiences, providing elements for improving strategy definition, increasing the efficiency and efficacy of policy intervention, and demonstrating the effects of intervention. All these elements help to improve the transparency and accountability demanded on the policy making process. As a rapidly growing and expanding field, the evaluation discipline is closely connected to the policy decision making process. However, there is still a lot of confusion on what evaluation can realistically and effectively achieve.

The theory and practice of evaluation has co-evolved with the developments experienced in science and technology policy. Evaluation tools have expanded to provide not only a means for quantification of policy impacts, but to facilitate mutual learning from past experiences, supporting mediation, decision-making and policy strategy definition. Increase in the complexity and uncertainty present in policy decision making requires the emergence of strategic intelligence combining the synergies of capacities between evaluation, technology foresight and technology assessment, to produce objective, politically unbiased, independent information to support active decision-making.

This toolbox of evaluation methods and approaches combines the effort of a network of evaluation experts to summarise the knowledge available regarding the main instruments applied in science and technology policy, emphasising the methodologies employed to evaluate their socio-economic impacts. Rationales for policy implementation, applied evaluation methodologies, conditions for application, methodology scope and limitations, data requirements and good practices are all addressed in this toolbox. The methodologies presented are not intended to be interpreted as close substitutes, as most likely, all evaluations will effectively benefit from a combined use of various of these methodologies. For instance, quantitative evaluation methods combined with performance indicators permit to capture the dynamics involved in science and technology providing good estimates of output and impact of public intervention. Policy makers could make use of these impact estimates as a means to legitimise and as supporting evidence of the rationale behind policy intervention. Qualitative evaluation methods provide policy makers with more detailed insights on the multiple effects of policy intervention and may help improving the processes and instruments of science and technology policies.

The toolbox is structured into five interrelated sections. The first section defines the user perspectives in evaluation, indicating the sometimes conflicting expectations arising between evaluation actors. The second section is structured around a set of broadly defined policy instrument categories providing the rationales for public intervention, main evaluation techniques available for their evaluation, and specifying illustrative practical examples of how evaluation have been conducted in the past. The third section introduces a series of evaluation methodologies, providing a description, reviewing the requirements for their application and illustrating their use with good practice examples. The fourth section analyses the conceptual framework describing how to incorporate evaluation practices within a "distributed techno-economic intelligence system" to inform and assist the policy making process. The fifth section gives some indications on how to use the capacities accumulated in the evaluation of the new policy instruments envisaged for supporting and promoting RTD activities.

The main content can be summarised as follows: Existing EU RTD evaluation practices (which comprise continuous monitoring, five year assessments and mid-term evaluation) are characterised by a strong focus on monitoring as compared to impact assessment, on projects and programmes rather than the broad context of policy and a heavy reliance on expert panels instead rather than on studies. Also, there is a constraint imposed by the limited time and monetary resources devoted to evaluation. With the increasing diversity of RTD policy instruments (e.g. funding of collaborative R&D, support to R&D infrastructure, measures for technology transfer and diffusion, standards and regulations, IPRs, networking…) relevant at the EU level, a sensible mix of available methodologies has to be applied in evaluations.

In this toolbox we give indications of the scope and limits of various methods. We find among others that the result of innovation surveys have to be better linked to the evaluation exercises on the macro level, the econometric impact assessments can be used on a wider scale as currently in use. This holds true not only for the macro-level but also – and especially – for micro-econometric tools like control group approaches. In the future, also cost-benefit analysis might play a larger role (in the arsenal of evaluators) not least in the ex-ante and ex-post evaluation of large-scale projects. Even with peer review – probably the most widely used approach in European S&T evaluation – it is possible to outline a potential for improved applicability, based on refined ways of panel composition, task allocation and decisory power.

But probably as important as the suggestions and recommendations with respect to individual tools and approaches is the perception that evaluation – to serve its purpose to empower policy learning – should follow some general good practice rules and ought to be embedded in a broader “system of distributed intelligence” which comprises other sorts of assessments of policy as well: benchmarking, foresight and technology assessment. It is in this setting that the use of evaluation results will yield most benefits to the policy matter.

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, more thorough evaluations at programme or policy level will require advances in knowledge both in the causal relations between inputs and outputs as well as to arrive at meaningful ways to measure and to aggregate these outputs.

Moreover, the continuous evolution of science and technology policies in modern economies requires devising methods to evaluate the efficiency and efficacy of new policies. For instance, the European Research Area (ERA) concept requires to better understand the increasing interconnection and integration between science and technology organizations and the rest of the involved actors. This would require implementing methods that allow to evaluate institutional capacity within changing environments. Thus, evaluation practices are increasingly required to analyse the effect science and technology policy induces on actors’ behaviors and institutional change.

All in all, the effort of synthesis presented in this RTD evaluation toolbox should be interpreted as a move towards introducing good governance practices in science and technology policy.

1 Guide to the Reader

Wolfgang Polt and Jaime Rojo

Evaluation is a systematic and objective process designed to assess the relevance, efficiency and effectiveness of policies, programmes and projects in attaining their originally stated objectives. It constitutes an eminently practice-driven approach, generating relevant information to feed back into the policy making process. Evaluation should be perceived as part of a continuous institutional learning process, bringing transparency and accountability to the decision making process and enabling to provide a justification for public intervention.

The main types of evaluation which might be conducted during the policy intervention cycle include:

- *Ex-ante or planning evaluation* is concerned with the design of the intervention, the definition of project objectives and how they will be achieved. The expected efficiency end effectiveness of the policy will need to be assessed.
- *Monitoring* is concerned with the ongoing collection and reviewing of information on the intervention implementation, its coverage and use. Monitoring is used to assess the quantity, quality and timeliness of the inputs, providing continuous feedback on the implementation of the policy intervention.
- *Intermediate evaluation* is concerned with reviewing the progress of the policy implementation. The results permit to adjust and improve targeting of the intervention.
- *Ex-post evaluation* is concerned with measuring the effects of the policy intervention. Evaluation provides information on the changes in the conditions and behaviour of targeted groups and individuals. The results might be used to improve design and implementation of the intervention.

The results of these evaluation processes are often complementary providing policy makers with relevant information on the design, implementation and effects of policy intervention.

The aim of the research and technological development (RTD) evaluation toolbox is to provide stakeholders concerned with evaluation issues, namely policy makers, programme managers and evaluation specialists with a synthetic overview on available policy instruments confronting them with the methodologies used in their evaluation. Aiming at improving the general understanding on the key elements conforming RTD policy evaluation, the concepts and the methodologies are presented in a rigorous and accessible way, illustrated providing multiple practical examples and giving significant emphasis to the assessment of the social and economic effects of policy intervention.

The compilation effort is the result of the work and experiences of the members of the EPUB network funded under the STRATA programme of the 5th Framework Programme.

Within its limited resources it could not strive for a complete coverage of all available policy instruments and evaluation methodologies. Rather, it attempts to provide some major building blocks on which future efforts in this vein can be based.

Especially with a view to the new challenges for evaluation practices and methods arising from the new Framework Programme and the European Research Area, this toolbox will be developed further.

Though presented as a ‘toolbox’, it cannot and should not be used mechanistically. If used to inform the stakeholders mentioned above about the scope and limits, the data requirements and some good practice examples that one has to consider in designing, commissioning or using an evaluation, it has served its purpose.

The toolbox is structured into five interrelated sections and a concluding section. The first section brings evaluation into context defining the user perspectives in evaluation, signalling the sometimes conflicting expectations arising between evaluators and policy makers. The second section is structured around a set of broadly defined policy instrument categories providing the rationales for public intervention, signalling the main evaluation techniques available for their evaluation, and specifying illustrative practical examples of past evaluations. The third section introduces a series of evaluation methodologies, providing a description, reviewing the requirements for their application and illustrating their use providing good practice examples. The fourth section analyses the conceptual framework describing how to incorporate evaluation practices within a “distributed techno-economic intelligence system” to inform and assist the policy making process. The fifth section gives some indications on how to use the capacities accumulated in the evaluation of the new policy instruments envisaged for supporting and promoting RTD activities. In the concluding section the coordinators offer a commented recap on the toolbox contents extracting a series of recommendations on good practices to consider when confronted with planning or conducting an evaluation.

2 User Perspectives

Mark Boden and Elliot Stern

Matching the requirements of policy makers with the skills and experience of evaluators can reveal crucial differences in perspectives. These may affect the delivery and implementation of evaluation studies that serve the desired policy purposes. These differences in perspective have been caricatured[1] as two gaps: the "delivery gap" between what policymakers want and what evaluators say, and the "customer gap" between what evaluators want and what policymakers say. While these caricatures combine and contrast features of ideal and real imagined situations, they set the scene for more serious consideration of matching user needs with available tools.

The Delivery Gap

The delivery gap illustrates what policy makers would ideally want from an evaluation to inform policy decisions and what evaluators believe is actually feasible. In the real world, the diffusion of knowledge is a complex process, governed by interactions between various knowledge producers and users. While policymakers may need information to inform spending decisions, evaluators might remind them that research may require years to have effects. While evaluators want clear attribution of effects to investment, a linear perspective on funding and output is usually unrealistic and additionality is complex to assess. Independent evidence of research excellence may also be unachievable given the loyalty of peers to their subject field and international colleagues. Also, while indicators to monitor and benchmark research performance are highly desirable, a crude regime may distort performance and be open to manipulation. Table 1 briefly indicates the main issues contributing to the delivery gap.

Table 1 The Delivery Gap

What policymakers want	What evaluators say
• Information in time for spending decision	• Research may take years to have effects
• Clear attribution of effects to investment	• Linear model is a rare case and additionality is complex to assess
• Independent evidence of research excellence	• Peers defend their subject field & international colleagues
• Key indicators to monitor & benchmark	• Crude regime distorts performance & can be manipulated

Customer Gap

In the converse situation, evaluators ideally need to have a clear, comprehensive and logical picture of the programmes they are evaluating, together with freedom and adequate resources. However, the real world of the policymakers is also complex and constrained. While evaluators want clearly defined and hierarchical programme objectives against which assess outcomes, policymakers may tell them that their programmes are a compromise involving multiple and conflicting objectives. While guaranteed independence might be desirable, this may also be compromised by the need for realistic recommendations within current policy constraints. Many evaluations do not have the luxury of time and resources for a full and thorough investigation, but face deadlines. Likewise, time constraints on programme management and participants are not conducive to giving evaluators full access to information and stakeholders.

[1]Georghiou, L. (2001) "The Impact and Utility of Evaluation", Conference on International best practices in evaluation of research in public institutes and universities, Brussels, 16.10.01

Figure 1 A Framework for a Process of Informed Design

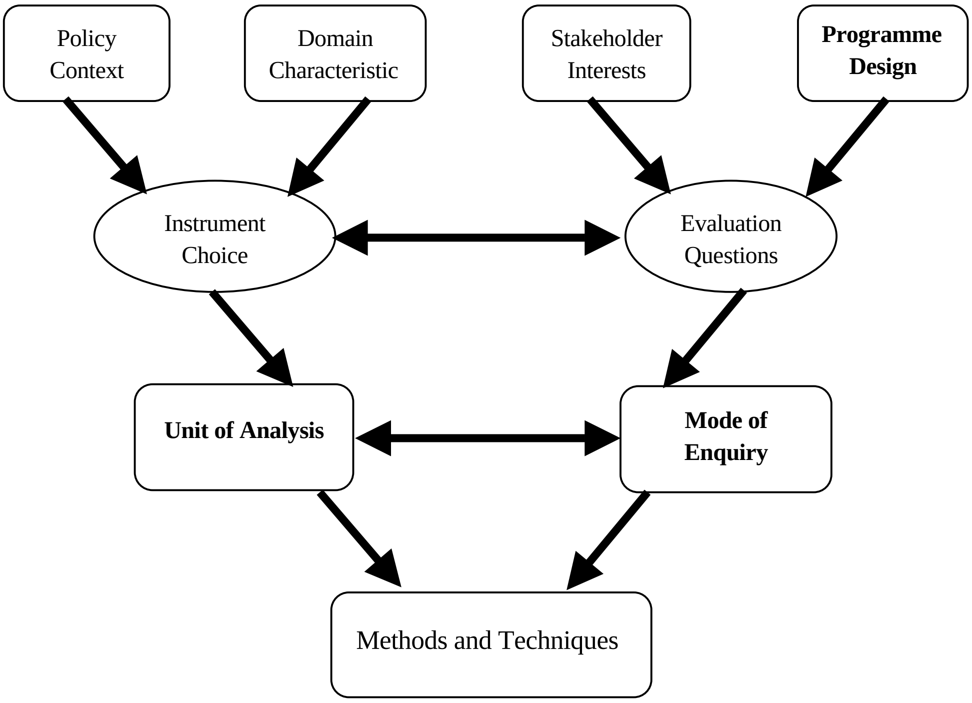

Table 2 The Customer Gap

What evaluators want	What policymakers say
• Clearly defined & hierarchical objectives	• Programmes are a compromise involving multiple & conflicting objectives
• Guaranteed independence	• Recommendations must be within realistic policy constraints
• Time & resources to do the job	• We need the results in three months
• Full access to information and stakeholders	• Everyone is overworked and busy

A longstanding weakness of many evaluations across many domains is their tendency to be method or technique led. The approach is determined by how rather than why or what questions. Particular experts with a commitment to their toolkit seek to apply their favoured approach. Any evaluation or analytic approach brings with it a certain logic, philosophical stance, range of possible outcomes and findings. New approaches to RTD policy evaluation need to be set into a framework that privileges questions of purpose, of ends rather than means. This introduction is intended to outline the kinds of questions evaluators and policy-makers should and do ask themselves when designing evaluations and as part of that design, before they come to select methods and techniques.

Figure 1suggests a framework for a process of informed design. By going through the process outlined in this figure policy makers should be better able to select appropriate evaluation instruments, fit for purpose.

The structure of this chapter attempts to reflect this framework, examining first the relationship between evaluation and the programme context in which it may operate. This takes into account the nature of the policy context, the characteristics of the scientific and technological domains that are relevant to it and how together these shape the choice and mode of implementation of policy instrument. This is followed by consideration of the design and implementation of the instrument, stakeholder issues and how these then relate to the evaluation questions that may be posed.

Against the background of the relationship between policy and evaluation, attention is turned to the organisation of evaluation, matching the unit of analysis selected and the mode of enquiry to the methods and techniques outlined in subsequent chapters/sections of the E-pub book/report.

The preparation of this chapter/report has drawn on a combination of personal and documentary sources. Any adequate account of user perspectives requires consultation with users, and this has entailed a series of interviews at European and national levels, with policy-makers, programme managers and other stakeholders.

On this basis this chapter offers insights into the design, use and impacts of evaluation at both national and European Commission levels, and, in particular, considers implications for the construction and use of a toolkit of evaluation methods.

3 Policy Instruments

3.1 INTRODUCTION

Wolfgang Polt and Jaime Rojo

There is nowadays a broadly shared perception of technological progress being the main contributor to long-term economic growth and consequently to the improvement of living standards and the quality of life. In this respect, R&D activities are one major source for producing the knowledge and ideas required to sustain technological progress. Theoretical and empirical models have indicated the relevance of technological progress in productivity growth. However, there is still need to improve the current understanding of the process of knowledge production, the connections between R&D, technology and diffusion, the magnitude of the contribution of technology to economic growth and the role of public intervention in the advance of technological progress.

3.1.1 Rationales for government support to R&D

Government support to R&D is generally justified on the grounds of existence of a market failure which leads to an underinvestment in R&D. The market failure appears in situations where firms would perform less R&D than is desirable from a social perspective, that is, in situations where due to imperfections in the market allocation mechanisms, the market will fail to allocate an efficient or socially optimal quantity of resources to the R&D. Theoretical and empirical research has shown the existence of a market failure in R&D provision. The main rationales for government support to R&D include:

- *Positive externalities or spillovers*: This results when the benefits from R&D activities are not fully captured by the R&D performer. Part of the benefits from performing R&D spillover to other individuals, firms or even across economies. This is due to the public good characteristic of knowledge and ideas, namely non-rivalry and non-excludability. In this respect, the empirical literature has found evidence on social rates of return to R&D exceeding the private ones and on the existence of significant spillovers to R&D.
- *Risk and uncertainty*: The large risk and uncertainty involved in R&D activities might deter private investment.
- *Network externalities*: The benefits of the technology increase with the pool of adopters.
- *Asymmetric information*: This situations appear in agent-principal relationships where one of the parts in the transaction the agent, has more information than the other, the principal.
- *Indivisibilities*: the large investments required to produce results might prevent firms from conducting R&D.
- *Evolutionary approaches*: technological progress benefits from enhanced competition and R&D diversity. Agents when making economic decisions are subject to bounded rationality which leads to the common pattern observed in R&D of path dependency and lock-in effects.

3.1.2 Mechanisms supporting R&D

Most countries have introduced mechanisms to support R&D to tackle the different types of market failures. Although the ultimate purpose of most of the instruments is to enhance technological progress, living standards and quality of life, the available instruments differ on their direct pursued objectives. The policy instruments described in the toolbox are framed within the context of the methods used for the evaluation and assessment of their impact. They have been grouped into four broad categories:

- *Financing R&D*: These interventions aim to compensate firms and individuals conducting R&D for the spillovers to society these activities generate and that they cannot appropriate. Main interventions in this area include direct subsidies where the government has to select the type of R&D conducted and indirect tax incentives where the firms select themselves the type the of R&D they want to conduct.
- *Provision of R&D infrastructure*: The government uses direct intervention in those situations where the market incentives are weak and spillover benefits are likely to be large and pervasive across sectors. The most commonly used measures include direct support to R&D infrastructures and provision of government sponsored R&D to PPPs in the formatation of infrastructures (e.d. networks of exellence).

- *Technology Transfer and Innovation Diffusion*: The policy intervention intends to increase social welfare by stimulating the diffusion of knowledge and the transformation of research results into commercial products. The most commonly used measures are schemes stimulating cooperation, RJV, science-industry collaborations, mobility of researchers, spin-offs, etc.
- *Legal Framework*: Interventions in this field pursue reducing existing market failures in the provision of private R&D. In the field of intellectual property rights (IPRs), intervention enhances the private incentives for conducting R&D by allowing firms to exert partial excludability for the use of the knowledge produced in the R&D process. The regulatory framework and the setting of standards also affect the path of technological development.

The purpose of Table 3 below is to assist the reader in the process of browsing through the contents of the RTD evaluation toolbox. It provides a simplified matrix matching the categories of public RTD policy instruments available with the methods adapted to evaluate their socio-economic impact.

Table 3 Evaluation Matrix: Matching policy instruments and methods

	Innovation Surveys	Econometric Models	Control Group Approaches	Cost Benefit Analysis	Expert Panels/ Peer Review	Field / Case Studies	Network Analysis	Foresight/ Technology Assessment	Benchmarking
Financing R&D	●●●	●●●	●●●	●		●●●		●	●
Provision of R&D infrastructure		●●		●●●	●●●	●●●	●●●	●●	●●●
Technology transfer/ innovation diffusion	●●●	●●●	●●●	●●	●	●●	●●●	●●●	●●●
Legal frameworks (IPRs, standards and regulation)	●	●	●	●●●		●●●		●●	●●●
Integrated projects			●	●●●	●●●	●●●	●●●	●●	●●
Networks of excellence					●●●	●●	●●●	●●	●●

Methodology: ●●● *relevant* ●●*sonewhat relevant* ●*low relevance*

3.1.3 Level of intervention

Although evaluation can be applied at any level of policy intervention, the following distinction is generally applied:

- Policy: a set of activities which may differ in type and may have different beneficiaries, directed towards common general objectives or goals.
- Programme: a set of organised but often varied activities bundled together –for example, projects, measures and processes– to achieve a common objective.
- Project: a single intervention with a fixed time schedule and dedicated budget.
- Thematic: is centred on a common objective pursued by several programmes.

3.1.4 Evaluation metrics and indicators

The objectives define the expected effects of the intervention. Once the objectives have been defined, indicators allow to evaluate the performance of the intervention and establish if the intervention is progressing towards meeting the defined objectives. An indicator is an objectively verifiable measurement which reflects the activity or effect being measured, allowing comparisons across different populations or individuals, and in time. Indicators are by definition imprecise measurements of the underlying concept of interest. The set of available indicators used in evaluation include:

- Input: are the resources consumed in the implementation of an intervention.
- Output: are the goods and services directly produced as a consequence of the intervention.
- Outcome: are the initial impacts of the intervention providing the reason for the programme. Outcomes tend to be less tangible than outputs.

- Impact: are the long-term socio-economic changes the intervention brings about.

Table 4 provides a useful illustration of the concepts and proxy variables available to capture the knowledge produced to measure the impact of R&D activities.

Table 4 Conceptual framework for knowledge measurement

Indicators	Concept	Proxies
Input	Persons-year, equipment and machinery-years	Expenditures
Output	Ideas, knowledge, invention	Publications, patents, prizes
Outcome/Impact	Advance of knowledge base	Papers, citations expert evaluation
	Societal improvement	Surveys and case studies, life expectancy
	Reduction of costs	Studies, statistical analysis costs and expenditures
	Economic output	Profit, revenues growth, revenue from new products
	Performance improvement	Productivity studies

Source: Adapted from Adam Jaffe (1998) Measuring Knowledge in the Health Sector, OECD/NSF High-Level Forum

The design of the policy intervention is the framework that permits to explain the expected achievements of the intervention and how the intervention its supposed to achieve its pursued objectives. In the evaluation process there is a need to establish how the inputs lead to the output and how these outputs subsequently lead to the outcome and impact that is expected from the intervention. One relevant aspect of the evaluation process is to identify and analyse the implicit assumptions and causal linkages behind the policy intervention.

3.1.5 Operational steps for conducting an evaluation

Ex-ante evaluation

The key aspects to assess the potential socio-economic benefit of policy intervention include:

- Establish a firm justification for public intervention, identifying the rationale and market failures addressed, and discussing how and why public intervention is appropriate.
- Introduce the modelling approach and the assumptions on which the projections are made, as results might be very sensitive to the modelling assumptions.
- Analyse the counterfactual based on constructed scenarios. What would happen with and without the project.
- Present cost-benefit or cost-effectiveness analysis comparing the various alternatives available. Evaluate alternatives of with project scenarios and with-out project scenarios. Provide indications based on the scenario projection on the costs considered and the benefits of the project.

Ex-post evaluation

The key aspects to conduct an ex-post evaluation of the socio-economic impact of a policy intervention include:

- Provide a clear specification of policy objectives.
- Define data collection at the programme design stage.
- Design the evaluation approach, with possible interaction with the evaluation sponsor.
- Control for the counterfactual, that is what would have occurred had the project not taken place.
- Compile evidence on success and failures.
- Provide clear results and recommendations when writing the evaluation report.

Evaluation design

Evaluation design allows evaluators to quantify the magnitude of the effects of a policy intervention while eliminating other competing explanations of the experimented change which are not related to the policy intervention. The two main alternatives in evaluation design are:

- *True experimental designs:* although appealing for the straightforward interpretation of the provided results they are rarely used in policy evaluation. This approach uses random selection to define programme participants (treatment group) and non participants (control group), but its implementations often finds resistance due to ethical, political or operational reasons.
- *Quasi experimental designs:* Quasi experimental designs are commonly used in evaluation and provide confident conclusions on the effect of the programme when the selection method for the comparison group is carefully implemented. The application of the method tries to approximate as much as possible the true experimental conditions. Besides their difference relating to participation in the programme, the characteristics of the comparison group should be as similar as possible to the treated group.

Evaluation implementation

The baseline data information is a key element in every evaluation. The baseline information requires to collect information referring to the situation before the policy intervention both on the participant and control (comparison) group. It will permit to measure with confidence the magnitude of change produced by the intervention. In simple before-after estimations, it allows to analyses the effect of the intervention by comparing post-intervention to pre-intervention status on selected indicators.

A relevant step in the preparatory procedure for conducting the evaluation is to establish the sample size and the sampling design method. The sample size is required because most projects are too large to permit the evaluation of the effects in all the participants. It will depend on the numbers of groups being studied, the amount of change expected in the indicator, confidence level required in the conclusions and the probability of detecting a difference in the indicators when one exists. The sampling method is required for selecting the individuals that will be used in the evaluation. The evaluation through the use of the control groups should also control for confounding factors, namely factors independent of the policy intervention that explain the changes observed in the supported individuals.

Data collection techniques

There are various methods for collecting data for evaluation purposes. Case studies and interviews, offer rich detailed data on the context in which an intervention is conducted and the problems it might solve. On the contrary, surveys provide quantitative data to measure the effects of the intervention. As reflected in Table 5 each technique has its own strengths and limitations. The adoption of appropriate approaches in data collection might reduce the data collection bias.

Table 5 Strengths and limitations in data collection

Data method	Strengths	Limitations
Interviews and case studies	quick implementation and low costs provides rich contextual information reveal project issues originally not thought	difficulty to code and analyse responses to open-ended questions difficulty to compare across interventions conduct of interviews requires expert staff
Surveys	capture information on inputs, outputs and impact allows to analyse broad range of issues information provided is easy to analyse possibility to generalise to the population if sufficient information	oversimplification of the process as a result of closed-ended questions interview bias expensive and time consuming

3.1.6 Policy instruments module structure

Each of the four broad categories of policy instruments described in the evaluation toolbox is structured to cover a common set of topics with the aim of facilitating comparisons across instruments.

- *General description*: The section provides a brief descriptive introduction to the policy category introducing the most relevant individual policy instruments grouped under this category. e.g. under the category of financing R&D describe briefly some of the most relevant policy instruments such as R&D subsidies, tax incentives.
- Policy objectives/outcomes: The section provides the objectives looked for by the policy instrument (e.g. competitiveness, quality of life, employment) and indicates the ways through the instrument can be evaluated against the stated objectives.
- Evaluation methodologies/tools applied and good practices examples: The section illustrates the methodologies and tools used to evaluate the impact of the instrument. The section is completed with a selection of good practice examples in the evaluation of the policy instrument at different levels (firm, industry, economy-wide) and for different fields (health consumer protection, environment, etc.).
- Data requirements/indicators: The section signals the data requirements needed for performing a monitoring and impact assessment of the instrument providing the main indicators used for measuring input and output.
- Illustration of operational approach to evaluation of instrument: The section provides a canonical illustrated example on how to perform an evaluation of the policy instrument.
- Commented bibliography for further insights: The section provides a commented bibliography for further insight on the evaluation of the instrument.

4 Evaluation Methodologies

4.1 INTRODUCTION

Wolfgang Polt and Jaime Rojo

The set of evaluation methodologies reviewed in the toolbox constitute recognised techniques particularly adapted to evaluate the socio-economic impact of RTD policies. Each of the evaluation methodologies presented has its own strengths and limitations, and the evaluator has to select them according to the particular requirements of the evaluation.

Evaluation methodologies have been adapted to capture the benefits of policies not directly measurable 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. Other techniques, such as hedonic prices, have been developed to adjust prices series to reflect changes in the quality of products, particularly relevant in high technology products subject to rapid quality changes.

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 permits to relate output measures with the broader impact of the policies on economic welfare. Most of the times, with the purpose of simplifying the complex dynamics of scientific and technological development, evaluation approaches adopt the operation of the linear model of innovation. 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, use more complex approaches such as national systems of innovation, chain linkage models of innovation and models of creative destruction.

The variables of interest in evaluation, outcome and impact, are captured indirectly by analysing how the outcomes of intervention change as a result to changes in the relationships and quantities of inputs and outputs. As evaluation approaches recognise, institutional variables are a fundamental factor shaping the effects of policy interventions. The different evaluation methodologies reviewed specially the qualitative ones indicate the relevance of the context dependence in evaluation.

Methodological approaches have been developed to deal with uncertainty and risk considerations for example, use of decision theory and expected return calculations in a cost benefit analysis contex. Uncertainty considerations can be introduced in ex-ante evaluation to provide a more accurate assessment of the return on R&D investment.

4.1.1 Ex-ante evaluation methodologies

The methodologies employed in ex-ante evaluation of programmes and policies include:

- *Foresight studies*: this structured consensus building methodology based on experts judgements permits to anticipate social, economical and technological development opportunities in policy planning.
- *Modelling and simulation*: this quantitative methodology uses scenario modelling to estimate the socio-economic impact of policy.
- *Cost-efficiency techniques*: this judgement methodology quantifies the costs and benefits associated with the policy intervention.
- *Cost-benefit techniques*: this judgement methodology compares in monetary terms all social and private cost and benefits of a programme to establish whether the benefits exceed the costs. The technique can be adapted to incorporate uncertainty and risk

Ex-ante economic evaluation methodologies can be a successful mechanism to:

- improve the performance of interventions
- allow funding in relation to merit
- ensure that the rationale of the funded interventions has been analysed
- analyse the expected benefit of an intervention

4.1.2 Monitoring and Ex-post evaluation methodologies

Monitoring and ex-post evaluation uses a combination of qualitative, statistical and econometric techniques to analyse the effects of the policy intervention. The diversity of methodologies available for performing an evaluation are a signal of the multiple dimensions in which the impacts of policy intervention might manifest themselves. For this reason, no single best evaluation methodology exists. Each methodology will be fitted to analyse particular dimensions of impacts, but the best evaluation approach would require a combination of various evaluation methodologies possibly applied at various level of data aggregation. Proceeding this way will allow cross checking the robustness on the observed effects of the intervention.

In the evaluation of the impact of RTD policies, there exists an added difficulty requiring to devise methodologies which allow to capture the effect produced by what is inherently an intangible good, the production and diffusion of knowledge. Most of the methodologies therefore focus on capturing particular relevant aspects of the RTD process. For instance, microeconometric methodologies allow to evaluate the existence of an additionality effect in public support on private R&D and to capture the private rate of return on R&D. Macroeconomic methodologies are better fitted to capture the generated R&D spillovers and longer term wider effects of the policy intervention on productivity and economic welfare. At a more disaggregated level, cost-benefit analysis, although not without difficulties, permits to transform all the benefits and cost of a project or programme intervention into monetary values and compare them. Network analysis and case studies are better fitted to capture the richness of impacts enabled in the RTD process, allowing to frame the evaluation in the broader socio-political context, and analyse accordingly the relevance of the social connections in the process. These methodologies are particularly relevant for the evaluation of institutions and systems.

The methodologies employed in ex-post evaluation of RTD programmes and policies include:

Statistical data analysis

- *Innovation Surveys*: provides basic data to describe the innovation process, summarised using descriptive statistics.
- *Benchmarking* allows to perform comparisons based on a relevant set of indicators across entities providing a reasoned explanation their values.

Modelling methodologies

- *Macroeconomic modelling and simulation approaches*: allows to estimate the broader socio-economic impact of policy interventions.
- *Microeconometric modelling*: permits to study the effect of policy intervention at the level of individuals or firms. There are mechanisms to control for the counterfactual by specifying a model which allows to estimate the effects on the outcome of the participant had the programme not taken place.
- *Productivity analysis*: permits to assess the impact of R&D on productivity growth at different levels data aggregation. This is particularly relevant to analyse the broader effects of R&D on the economy.
- *Control group approaches*: allows to capture the effect of the programme on participants using statistical sophisticated techniques.

Qualitative and semiquantitative methodologies

- *Interviews and case studies*: uses direct observation of naturally occurring events to investigate behaviours in their indigenous social setting.

- *Cost-benefit analysis*: allows to establish whether a programme or project is economically efficient by appraising all its economic and social effects.
- *Expert Panels/Peer Review*:measures scientific output relying on the perception scientists have of the scientific contributions made by other peers. Peer review is the most widely used method for the evaluation of the output of scientific research.
- *Network Analysis*: allows to analyse the structure of cooperation relationships and the consequences for individuals decisions' on actions providing explanations for the observed behaviours by analysing their social connections into networks.
- *Foresight/ Technology Assessment*: used to identify potential mismatches in the strategic efficiency of projects and programmes.
-

Table 6 and Table 7 attempt to briefly summarise the wealth of information contained in each of the evaluation methodologies modules. For a deeper appreciation on the methodology of interest it is advised to refer directly to the information provided in the detailed modules.

Ex-post economic evaluation methodologies have proved a successful mechanism to:

- determine the efficiency and efficacy of the intervention (e.g. macro economic simulation and productivity studies).
- provide a quantitative estimation of the impact of the intervention (e.g. microeconomic evaluation studies).
- quantify the various dimensions in which returns should be considered within a defined framework.
- assess environmental sustainability and health issues (e.g. cost-benefit analysis), organizational impact (e.g. case studies, network analysis, innovation studies), strategic impact (e.g. foresight).

4.1.3 Evaluation methodologies module structure

Each of the eleven evaluation methodologies described in the evaluation toolbox is structured to cover a common set of topics with the aim of facilitating comparisons across methodologies.

- *General description*: The section provides a general description of the evaluation methodology.
- *Policy instruments/interventions to evaluate with the method*: The section discusses the policy instruments evaluated with the methodology.
- *Good practice examples*: The section illustrates with practical examples how the methodology as been used in RTD evaluation at different levels (firm, industry, aggregated). Examples of good practices of method application.
- *Conditions for methodology application*: The section provides indications on the costs of implementation, complexity of use, phase of application (i.e. ex ante, intermediate, ex-post evaluation), degree of methodology acceptance.
- *Operational steps for method application*: The section provides an illustrated example on how the method is operationally used in evaluation and comments the relevant concepts generally addressed (e.g. additionality, spillovers, cluster formation, rate of return, etc) and on particularly relevant methodological aspects to consider (e.g. panel data, non linearities, time series)
- *Data requirements/indicators*: The section provides an indication of the input, output, outcome/impact indicators available to measure the output and the outcome of the policy instrument.
- *General assessment of the scope and limits of methodology*: The section discusses the scope and limitations of the methodology use in evaluation.
- *Commented bibliography for further insights*: The section provides references for obtaining further insights on the methodology.

Table 6 Evaluation Methodologies

Methodology	**Quantitive /Qualitative Ex-ante Monitoring/ Ex-post**	**Data Requirements**	**Strengths**	**Limitations**	**Good Practices**
Innovation Surveys	Semi-quantitative Quantitative *Monitoring* *Ex-post*	Micro data Expenditures Profits Patents, Innovation	Detect innovation trends and insights on the soft side of innovation. Findings from interviewed sample can be generalised to the population Permits to identify size and distribution of impacts Provides groups comparisons and changes over time	High cost and time consuming Processing and analysis of data requires large human resources Some types of information are difficult to obtain Long time series generally not available	Analysis of the innovation process using data on the EU Community Innovation Survey
Micro Methods	Quantitative qualitative categorical data *Monitoring* *Ex-post*	Micro data Expenditures Profits Patents	Results based on explicit formulation of theory based causal relationships R&D Additionality Control for different effects: firm size, expenditures, innovation capacity	Quality of data Persuade participant and non participant entities to disclose information Only private rate of return to R&D	Effects of public R&D subsidies on firms in West Germany Evaluation of the ITF Programme FlexCIM Effects of R&D subsidies in Spain Promotion of AMT technologies based on Swiss Micro data
Macro Methods	Quantitative modelling methodology *Ex-ante (simulation)* *Monitoring* *Ex-post*	R&D Expenditures R&D output Macroeconomic data	Social Rate of return to R&D Capture R&D Spillovers Estimate long term policy intervention impact Scenario simulations for policy supported geographical areas	Average returns Robustness of results Time lags for observation of the effects	Modelling approaches: OECD Interlink, IMF Multimod, EU Quest. R&D Spillover studies: Jaffe, Nadiri International spillovers: Eaton and Kortum, Mohnen, Evenson
Productivity Studies	Quantitative modelling methodology *Monitoring* *Ex-post*	Micro data Expenditures Profits R&D, Patents	Estimation of effect of R&D on productivity Estimate the rate of return to R&D	Quality of data Deflation of series Required assumptions for measurement of stock variables	Productivity studies (Van Ark) Growth accounting (Griliches, Jorgenson) Micro datasets: French INSEE and US Census of Manufacturers
Control group approaches	Quantitative *Ex-post*	Micro data Expenditures Profits Patents	Capture the impact of policy intervention on the programme participant entitity	Requires high technical capacity High Implementation Cost Data Demanding	Collaborative industrial Research between Japan and US Evaluation of RTDI instruments in Ireland Participation of Ireland in Eurpoean Space Agency
Cost Benefit Analysis	Quantitative (with qualitative elements) *Ex-ante (especially)* *Monitoring* *Ex-post*	Micro data Profit & cost estimates	Provides an estimate of socio-economic effect of intervention Good approach to assess the efficiency of an intervention Addresses by making them explicit all the economic assumptions of the impact of the intervention	Requires high technical capacity Some degree of judgement and subjectivity, depends on largely on assumptions made Not easily comparable across cases Careful interpretation of results when benefits are not easily quantifiable in monetary terms	US Advanced Technology Programme US National Institute of Standards Methodology
Expert Panels /Peer Review	Qualitative Semi-quantitative *Ex-ante* *Monitoring* *Ex-post*	Project programme data	Evaluation of scientific merits Flexibility Wide scope of application Fairness	Peers independence Economic benefits not captured	Evaluation of Large Infrastructures Evaluation of EU Programmes
Field /Case studies	Qualitative Semi-quantitative *Monitoring* *Ex-post*	Project programme data	Observation of the socio-economic impacts of intervention under naturalistic conditions Good as exploratory and descriptive means of investigation Good for understanding how contexts affect and shape impacts	Results not generalisable	Telematic innovation in the health care sector. Evaluation case studies reviewed in Georghiou and Roessner (2000)
Network Analysis	Qualitative Semi-quantitative *Ex-post*	Project programme data	Comprehensive empirical material compilation for policy purposes Cooperation linkages	Time involved in collecting the survey information Persuasion requirements	RTO systems Interdisciplinary centers of medical research
Foresight/ Technology Assessment	Qualitative Semi-quantitative *Ex-ante* *Monitoring*	Qualitative data Scenario	Consensus building to reduce uncertainty under different scenarios Combination on public domain and private domain data Articulation and road mapping of development of new technologies	Impossibility to detect major RTD breakthroughs	Benchmarking of ISI/FhG capacities against Foresight results
Benchmarking	Semi-quantitative *Ex-post* *Monitoring*	Science and technology Indicators	Comparison method across different sectors	Data detail requirements Non transferable	EU Benchmarking national policies Innovation Trend Chart

Table 7 Evaluation Methodologies II

Methodology	Data application level	Areas of application	Output	Outcome	Impact
Innovation Surveys	Firm Industry Economy-wide	Innovation IPRs Technology transfer Research collaboration	New products and processes Increase in sales Increase in value added Patent counts, IPRs	Creation of new jobs Innnovation capacity building	Enhanced Competitiveness Institutional and organisational efficiency, Faster diffusion of Innovation Employment
Micro Methods	Plant Firm Industry Economy-wide	Sectoral Returns to R&D	Output and value added (collect baseline info for before-after comparisons)	Sectoral productivity industry sectoral spillovers Additionality, Leverage effects	Firms competitiveness
Macro Methods	Firm Industry Economy-wide	Sectoral Regional Economy-wide	Output and value added	Change in R&D Capital, Human capital, Social capital International R&D Spillovers	Regional, country productivity Employment, Good governance Economic and social cohesion
Productivity Studies	Firm Industry Regional Economy-wide	Sectoral Regional Economy-wide	Output and value added	knowledge, geographical and International R&D Spillovers	Regional, country productivity Employment Economic and social cohesion
Control Group Approaches	Firm Industry	Technology implementation Innovation	Output and value added (on supported and non supported firma)	Additionality Rate of return to R&D	Firm, industrial competitiveness
Cost Benefit Analysis	Firm Industry	Health Environment Energy Transport	Value added benefit-cost ratio IRR Consumer surplus	Health improvements Consumer protection Environmental sustainability	Quality of life Standard of living
Expert Panels/ Peer Review	Firm Industry Economy-wide	Scientific merit Technological capacity	Publication counts Technological output	Scientific and Technological capabilities	R&D performance
Field/ Case Studies	Firm Industry	Science-industry relationships	Detailed inputs and outputs	firms RTD capabilities on the job-training educational schemes	Industrial competitiveness Quality of life Organisational efficiency
Network Analysis	Firm Industry Regional	RJVs, cooperation science industry Clusters	Cooperation linkages	Cooperation in clusters Social embededness	Efficiency of institutional relationships
Foresight/ Technology Assessment	Institution Regional Economy-wide	Technology Trends	Identification of generic technologies Date of implementation	Technological capacities	Technological paradigms shifts
Benchmarking	Firm Industry Economy-wide	Efficiency of technology policy	S&T indicators	Technology capabilities	Industry competitiveness Good governance

5 Distributed Techno-Economic Intelligence for Policymaking

Stefan Kuhlmann

5.1 INTRODUCTION

Who is in need of evaluations of public research, technology and innovation policies? And why, for which purpose? Meta-evaluations of evaluation practices give evidence of an increasing production of evaluative information for public policymaking in the area of research and innovation (e.g. Georghiou 1995; Kuhlmann/Holland 1995; Kuhlmann/Heinze 2001). At the same time, experience shows that both the theory and practice of evaluation has undergone important developments over the past decade. In particular, in countries where evaluation has taken root fairly early, the following *trends* can be observed:

- The major rationale for evaluations has shifted and evolved from an attempt to *legitimate past actions* and demonstrate accountability, to the need to improve understanding and *inform future initiatives.*
- Correspondingly, the issue focus of evaluations has broadened away from a narrow focus on quality, economy, efficiency and effectiveness of a programme, and towards *a more all-embracing concern* with additional issues, such as the appropriateness of a policy tool and a concern with performance improvement and strategy development.
- Approaches to evaluation have evolved from a purist model of "objective neutrality", characterised by independent evaluators producing evaluation outputs containing evidence and argument, but no recommendations; to more formative approaches in which evaluators act as process consultants and mediators in learning exercises *involving all relevant stakeholders*, providing advice and recommendations as well as independent analysis.
- This has led to more *flexible and experimental concepts* of policy portfolios, and to even greater demands for well specified systems of monitoring, evaluation and benchmarking to aid analyses and feed back into strategy development.

Many evaluation exercises thus reflect an increasing concern with the link between evaluation and strategy, with an varying mix of methodologies used within the context of individual exercises to satisfy the demands for understanding and advice. Increasing attention is also being paid within many institutional settings to the way in which *evaluation can inform strategy* – and quite often also in combination with benchmarking studies, technology foresight exercises, technology assessment efforts and other analytical tools. The combined use of such tools has been hallmarked *"strategic intelligence"*[2].

The present chapter discusses evaluation and neighbouring exercises as intelligence tools for research and innovation policies, in the context of 'systems of innovation' and related – often contested – arenas for policymaking (section 5.2). Some illustrative examples of strategic intelligence use will be sketched (section 5.3). Critically, also the need for a system of "distributed intelligence" is examined which could provide public and corporate policymakers with access to strategic intelligence outputs produced in different locations for different reasons. Specifically, the design requirements of a "system architecture" for distributed intelligence is explored (section 5.4). In section 5.5 briefly the need for an effective European system of distributed intelligence is contemplated. Finally, section 5.6 summarizes the guiding principles and related requirements.

---

[2] See Kuhlmann et al., 1999: the Advanced Science and Technology Policy Planning Network (ASTPP), a thematic network set up as part of the Targeted Socio-Economic Research (TSER) Programme of the European Commission developed an outline of 'Distributed Strategic Intelligence', providing a conceptual basis for the present chapter. The present text draws also on a related, more recent publication (Kuhlmann 2001a).

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

Luke Georghiou and Stefan Kuhlmann

6.1 INTRODUCTION

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

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

7 Concluding Remarks

Wolfgang Polt, Jaime Rojo

All modern economies adopt policies to support their science and technology system aiming to improve economic performance and social welfare. This increasing relevance of science and technology in shaping the forces driving our societies and economies is forcefully reflected in the Presidency's conclusions of the European Council held in Lisbon on March 2000 where the Heads of Government and State agreed on the new strategic goal of the EU to become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable growth with more and better jobs and greater social cohesion. Fulfilling this challenging strategy requires a clear commitment on the part of governments to discern the socio-economic impacts of science and technology policies.

Evaluation is a key decision support tool providing policy makers with a better understanding of policy results, allowing learning from past experiences, providing elements for improving strategy definition, increasing the efficiency and efficacy of policy intervention, and demonstrating the effects of intervention. All these elements help to improve the transparency and accountability demanded on the policy making process. As a rapidly growing and expanding field, the evaluation discipline is closely connected to the policy decision making process. However, there is still a lot of confusion on what evaluation can realistically and effectively achieve.

The theory and practice of evaluation has co-evolved with the developments experienced in science and technology policy. Evaluation tools have expanded to provide not only a means for quantification of policy impacts, but to facilitate mutual learning from past experiences, supporting mediation, decision-making and policy strategy definition. Increase in the complexity and uncertainty present in policy decision making requires the emergence of strategic intelligence combining the synergies of capacities between evaluation, technology foresight and technology assessment, to produce objective, politically unbiased, independent information to support active decision-making.

This toolbox of evaluation methods and approaches combines the effort of a network of evaluation experts to summarise the knowledge available regarding the main instruments applied in science and technology policy, emphasising the methodologies employed to evaluate their socio-economic impacts. Rationales for policy implementation, applied evaluation methodologies, conditions for application, methodology scope and limitations, data requirements and good practices are all addressed in this toolbox. The methodologies presented are not intended to be interpreted as close substitutes, as most likely, all evaluations will effectively benefit from a combined use of various of these methodologies. For instance, quantitative evaluation methods combined with performance indicators permit to capture the dynamics involved in science and technology providing good estimates of output and impact of public intervention. Policy makers could make use of these impact estimates as a means to legitimise and as supporting evidence of the rationale behind policy intervention. Qualitative evaluation methods provide policy makers with more detailed insights on the multiple effects of policy intervention and may help improving the processes and instruments of science and technology policies.

The toolbox is structured into five interrelated sections. The first section defines the user perspectives in evaluation, indicating the sometimes conflicting expectations arising between evaluation actors. The second section is structured around a set of broadly defined policy instrument categories providing the rationales for public intervention, main evaluation techniques available for their evaluation, and specifying illustrative practical examples of how evaluation have been conducted in the past. The third section introduces a series of evaluation methodologies, providing a description, reviewing the requirements for their application and illustrating their use with good practice examples. The fourth section analyses the conceptual framework describing how to incorporate evaluation practices within a "distributed techno-economic intelligence system" to inform and assist the policy making process. The fifth section gives some indications on how to use the capacities accumulated in the evaluation of the new policy instruments envisaged for supporting and promoting RTD activities.

The main content can be summarised as follows: Existing EU RTD evaluation practices (which comprise continuous monitoring, five year assessments and mid-term evaluation) are characterised by a strong focus on monitoring as compared to impact assessment, on projects and programmes rather than the broad context of policy and a heavy reliance on expert panels instead rather than on studies. Also, there is a constraint imposed by the limited time and monetary resources devoted to evaluation. With the increasing diversity of

RTD policy instruments (e.g. funding of collaborative R&D, support to R&D infrastructure, measures for technology transfer and diffusion, standards and regulations, IPRs, networking…) relevant at the EU level, a sensible mix of available methodologies has to be applied in evaluations.

In this toolbox we give indications of the scope and limits of various methods.

Evaluation Matrix: Matching policy instruments and methods

	Innovation Surveys	Econometric Models	Control Group Approaches	Cost Benefit Analysis	Expert Panels/ Peer Review	Field / Case Studies	Network Analysis	Foresight/ Technology Assessment	Benchmarking
Financing R&D	●●●	●●●	●●●	●		●●●		●	●
Provision of R&D infrastructure		●●		●●●	●●●	●●●	●●●	●●	●●●
Technology transfer and innovation diffusion	●●●	●●●	●●●	●●	●	●●	●●●	●●●	●●●
Legal frameworks (IPRs, standards and regulation)	●	●	●	●●●		●●●		●●	●●●
Integrated projects			●	●●●	●●●	●●●	●●●	●●	●●
Networks of excellence					●●●	●●	●●●	●●	●●

Methodology: ●●● *relevant* ●● *somewhat relevant* ● *low relevance*

As can be seen, there is hardly any instrument that would not be best evaluated using a number of different methods.

We find among others that the result of innovation surveys have to be better linked to the evaluation exercises on the macro level, the econometric impact assessments can be used on a wider scale as currently in use. This holds true not only for the macro-level but also – and especially – for micro-econometric tools like control group approaches. In the future, also cost-benefit analysis might play a larger role (in the arsenal of evaluators) not least in the ex-ante and ex-post evaluation of large-scale projects. Even with peer review – probably the most widely used approach in European S&T evaluation – it is possible to outline a potential for improved applicability, based on refined ways of panel composition, task allocation and decisional power.

But probably as important as the suggestions and recommendations with respect to individual tools and approaches is the perception that evaluation – to serve its purpose to empower policy learning – should follow some general good practice rules and ought to be embedded in a broader “system of distributed intelligence” which comprises other sorts of assessments of policy as well: benchmarking, foresight and technology assessment. It is in this setting that the use of evaluation results will yield most benefits to the policy matter.

Evaluation Methodologies

Methodology	Quantitive /Qualitative Ex-ante Monitoring/ Ex-post	Data Requirements	Strengths	Limitations	Good Practices
Innovation Surveys	Semi-quantitative Quantitative *Monitoring* *Ex-post*	Micro data Expenditures Profits Patents, Innovation	Detect innovation trends and insights on the soft side of innovation. Findings from interviewed sample can be generalised to the population Permits to identify size and distribution of impacts Provides groups comparisons and changes over time	High cost and time consuming Processing and analysis of data requires large human resources Some types of information are difficult to obtain Long time series generally not available	Analysis of the innovation process using data on the EU Community Innovation Survey
Micro Methods	Quantitative qualitative categorical data *Monitoring* *Ex-post*	Micro data Expenditures Profits Patents	Results based on explicit formulation of theory based causal relationships R&D Additionality Control for different effects: firm size, expenditures, innovation capacity	Quality of data Persuade participant and non participant entities to disclose information Only private rate of return to R&D	Effects of public R&D subsidies on firms in West Germany Evaluation of the ITF Programme FlexCIM Effects of R&D subsidies in Spain Promotion of AMT technologies based on Swiss Micro data
Macro Methods	Quantitative modelling methodology *Ex-ante (simulation)* *Monitoring* *Ex-post*	R&D Expenditures R&D output Macroeconomic data	Social Rate of return to R&D Capture R&D Spillovers Estimate long term policy intervention impact Scenario simulations for policy supported geographical areas	Average returns Robustness of results Time lags for observation of the effects	Modelling approaches: OECD Interlink, IMF Multimod, EU Quest. R&D Spillover studies: Jaffe, Nadiri International spillovers: Eaton and Kortum, Mohnen, Evenson
Productivity Studies	Quantitative modelling methodology *Monitoring* *Ex-post*	Micro data Expenditures Profits R&D, Patents	Estimation of effect of R&D on productivity Estimate the rate of return to R&D	Quality of data Deflation of series Required assumptions for measurement of stock variables	Productivity studies (Van Ark) Growth accounting (Griliches, Jorgenson) Micro datasets: French INSEE and US Census of Manufacturers
Control group approaches	Quantitative *Ex-post*	Micro data Expenditures Profits Patents	Capture the impact of policy intervention on the programme participant entity	Requires high technical capacity High Implementation Cost Data Demanding	Collaborative industrial Research between Japan and US Evaluation of RTDI instruments in Ireland Participation of Ireland in Eurpoean Space Agency
Cost Benefit Analysis	Quantitative (with qualitative elements) *Ex-ante (especially)* *Monitoring* *Ex-post*	Micro data Profit & cost estimates	Provides an estimate of socio-economic effect of intervention Good approach to assess the efficiency of an intervention Addresses by making them explicit all the economic assumptions of the impact of the intervention	Requires high technical capacity Some degree of judgement and subjectivity, depends on largely on assumptions made Not easily comparable across cases Careful interpretation of results when benefits are not easily quantifiable in monetary terms	US Advanced Technology Programme US National Institute of Standards Methodology
Expert Panels /Peer Review	Qualitative Semi-quantitative *Ex-ante* *Monitoring* *Ex-post*	Project programme data	Evaluation of scientific merits Flexibility Wide scope of application Fairness	Peers independence Economic benefits not captured	Evaluation of Large Infrastructures Evaluation of EU Programmes
Field /Case studies	Qualitative Semi-quantitative *Monitoring* *Ex-post*	Project programme data	Observation of the socio-economic impacts of intervention under naturalistic conditions Good as exploratory and descriptive means of investigation Good for understanding how contexts affect and shape impacts	Results not generalisable	Telematic innovation in the health care sector. Evaluation case studies reviewed in Georghiou and Roessner (2000)
Network Analysis	Qualitative Semi-quantitative *Ex-post*	Project programme data	Comprehensive empirical material compilation for policy purposes Cooperation linkages	Time involved in collecting the survey information Persuasion requirements	RTO systems Interdisciplinary centers of medical research
Foresight/ Technology Assessment	Qualitative Semi-quantitative *Ex-ante* *Monitoring*	Qualitative data Scenario	Consensus building to reduce uncertainty under different scenarios Combination on public domain and private domain data Articulation and road mapping of development of new technologies	Impossibility to detect major RTD breakthroughs	Benchmarking of ISI/FhG capacities against Foresight results
Benchmarking	Semi-quantitative *Ex-post* *Monitoring*	Science and technology Indicators	Comparison method across different sectors	Data detail requirements Non transferable	EU Benchmarking national policies Innovation Trend Chart

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, more thorough evaluations at programme or policy level will require advances in knowledge both in the causal relations between inputs and outputs as well as to arrive at meaningful ways to measure and to aggregate these outputs.

Moreover, the continuous evolution of science and technology policies in modern economies requires devising methods to evaluate the efficiency and efficacy of new policies. For instance, the European Research Area (ERA) concept requires to better understand the increasing interconnection and integration between science and technology organizations and the rest of the involved actors. This would require implementing methods that allow to evaluate institutional capacity within changing environments. Thus, evaluation practices are increasingly required to analyse the effect science and technology policy induces on actors' behaviors and institutional change.

With respect to good practice rules, the following can be synthesised from the information contained in the toolbox and from the network members' discussions in the context of this project:

Evaluation planning

- Introduce new methods in ex-ante evaluation that favour diversity and the taking up of new risks and multidisciplinarity. Peer review is a significantly conservative approach in the evaluation of research proposals, risky projects are likely to get worse scores from peer review. Mainstream science is better positioned when adopting peer review methods.
- Provide an early and adequate scheme for the evaluation design and integrate it into the policy intervention design to ensure that intervention objectives are clearly defined and can be effectively evaluated.
- Base the public intervention on a demonstrated market or systemic failure which the intervention should solve.

Operational and management issues

- Allocate sufficient time and monetary resources to evaluation. This is justified as the aim is to ensure that public money is efficiently and wisely spent.
- Promote independence to ensure credibility of results, for this purpose it might be relevant to use external evaluation experts (from other nation).
- Involve policy makers and project managers in the evaluation so that their views are fed into the evaluation and they contribute to the design of the evaluation approach.
- Separate in evaluation the strategy function from the operational function. Evaluation as a demonstration of impact is only one input to strategy definition.
- Strengthen transparency by publishing the terms of reference, criteria's used in the evaluation and disseminating the produced evaluation results to a broad audience of interested bodies.

Evaluation priors

- Clarify the implicit policy rationale of the intervention as a first step in conducting an evaluation.
- Identify and clarify the objectives of the policy intervention being evaluated in the context of the intervention logical framework including the implicit assumptions and establishing the feasibility of evaluating them.
- Define the requirements for data compilation and regular updating during the intervention design stage as ex-post evaluation largely depends on the quality of the compiled data.
- Define the intervention jointly with concrete targets that will facilitate the evaluation of the instrument, e.g. "increase the publications in the field of genetic technology by 20 per cent or increase productivity by 10 per cent".

- Ensure the compilation of data before and after the intervention as well as on supported and non supported units to allow to control for the counterfactual.

Method implementation

- Adapt methodologies to deal with the particular evaluation requirements and to answer relevant questions. Evaluation should not be perceived as mechanical process. Definition of objectives determine the methodology selection.
- Combine different methodologies and different levels of data aggregation to improve the understanding of the multidimensional effects of the policy intervention.
- Incorporate systemic considerations into evaluation as science and technology is likely to modify institutions structure and behaviour.
- Separate when possible the evaluation of the scientific merit provided by traditional established methods such as peer review from the evaluation of the other socio-economic objectives using the support of expert panels or peer review.
- Evaluate the profile of supported and non-supported firms including those rejected and those who did not apply for support. Control group approaches are especially valuable in this context.
- Establish intended and unintended effects of the intervention. Analyse failure as well as success histories.

Strategic evaluation

- Integrate evaluation practices with other sources of "distributed intelligence" such as technology assessment and foresight to support strategy policy definition.
- Develop new applications of benchmarking, foresight and network approaches to evaluate new concepts such as institutional capacity, behavioral additionality, networking.

Dissemination of evaluation results and recommendations

- Broaden the use of evaluation results by incorporating the views of the potentially interested audience such as industry, target groups and social communities representatives.
- Introduce the requirement that programme managers report on implementation of recommendations made in the evaluations.
- Produce timely evaluation reports and in a clearly and understandable language to increase impact.

All in all, the effort of synthesis presented in this RTD evaluation toolbox should be interpreted as a move towards introducing good governance practices in science and technology policy.

InTeReg Research Report Series

The Research Report Series seeks to disseminate the results of research conducted within the Institute of Technology and Regional Policy of Joanneum Research to the broad academic community and other interested parties. Since much of the research is ongoing, the authors welcome comments from readers.

Electronic copies of the Research Report Series can be found at: http://www.joanneum.at/rtg/rp.

For further questions, please contact intereg@joanneum.at.

JOANNEUM RESEARCH Forschungsgesellschaft mbH – Institute of Technology and Regional Policy (InTeReg)

Vienna Office:
Wiedner Hauptstraße 76
A-1040 Vienna, Austria
Phone: +43-1-581 75 20
E-Mail: intereg@joanneum.at

Graz Office:
Elisabethstraße 20
A-8010 Graz, Austria
Phone.: +43-316-876 1488
E-Mail: intereg@joanneum.at