



# **Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes**

**Report to the European Commission  
DG XII, Evaluation Unit**

by the

**Independent Reflection Group:**

Angelo Airaghi (Chairman)

Niels E. Busch

Luke Georghiou

Stefan Kuhlmann (Rapporteur)

Marc J. Ledoux

Anthony F.J. van Raan

José Viana Baptista

January, 1999

## **Contents**

<b>Summary</b>	<b>1</b>
<b>1. The need for assessment</b>	<b>4</b>
<b>2. Assessment: context, issues, definitions</b>	<b>6</b>
<b>3. Measuring impacts, relevance and effectiveness</b>	<b>9</b>
3.1 The nature of impacts of publicly-funded RTD	9
3.2 Methods and techniques for evaluation	12
3.3 Focus on economic and social impacts	13
<b>4. Focusing on the Framework Programmes: challenges and limitations</b>	<b>16</b>
4.1 Special features of the Framework Programmes	16
4.2 Widening the scope: the Fifth Framework Programme	17
4.3 Consequences for impact assessments	17
4.4 The time problem: overlaps between Framework Programmes	19
<b>5. Assessing future Framework Programmes: the way forward</b>	<b>22</b>
5.1 Basic requirements: feasibility, intelligence and debate	22
5.2 Discussing a basic European assessment model	23
5.3 RTD assessment in a broader European context	29
<b>Annexes</b>	<b>30</b>

## Summary

Europe's investment in RTD creates a demand for information on the efficiency with which RTD is managed, the quality of the work itself, and the economic and social returns. Evaluation schemes set up to supply this information are important tools for policymakers, and they give the research community an opportunity to demonstrate its achievements. This report looks at how we can improve our assessment of the impact of the RTD work supported by the EU.

The evaluation scheme introduced for the Fourth European RTD Framework Programme had two main components, continuous monitoring and five-year assessments. The Fifth Framework Programme is now placing additional demands on European RTD, especially in its socio-economic objectives and the need for more flexibility in policy matters. This in turn means that the evaluation scheme will have to be extended.

Despite the attractions of an investment view of RTD, it is widely accepted that normal rate-of-return criteria do not apply to RTD. There are three main problems:

- **RTD has many effects.** Short-term benefits to participants in RTD projects include increased sales and market share, improved skills and know-how, and new contacts, but the effects of RTD go well beyond this. A typical RTD project affects many organisations which do not participate directly in it, yet these impacts are hard to measure.
- **Evaluation results are often needed early,** sometimes even before the RTD work is complete and usually before its full socio-economic effect can be measured.
- **Effects are not always uniquely attributable to a single project or participation.** They may arise from a combination of inputs, both within and outside RTD, including the mechanisms for the exploitation of their results

The Fifth Framework Programme brings some additional challenges. Evaluations will henceforth need to pay more attention to the "European value added", including the contribution of each programme to the development of European "critical mass" in its subject area, contributions to Community policies and problems arising at Community level. In fact, the Fifth Framework Programme comprises so many diverse elements that no single methodological approach to evaluation will work for every area. Instead we need an evaluation system that respects the full diversity of Europe's RTD activities.

The special emphasis of the Fifth Framework Programme on social objectives also brings new challenges. We need to deal with a broader range of stakeholders and to measure the effects of RTD on employment, health, quality of life and the environment. However, most social change is driven by factors other than RTD policy, for instance by immigration, taxation and education policies.

For this reason it is normally impossible to attribute job creation and other specific social effects to a given RTD initiative. Moreover these RTD effects are hard to measure in financial terms. New evaluation methods need to look less at money and more at how laws, standards and practices change as a result of RTD. Just because its monetary benefits are hard to quantify does not mean that a project is not worthwhile.

An issue which needs to be addressed is the extent to which public funding has made a difference to a firm's R&D decisions. This 'additionality' question at its simplest explores whether the firm would have done the project in the absence of funding. The aim is of course to cause firms to increase their R&D spending, for example by undertaking projects with a higher risk but potentially higher payoff. Additionality can also be manifested by changing the way in which the project is done, for example making it larger, faster or more collaborative.

By stressing the limitations of impact evaluations we do not imply that these evaluations are not worthwhile. We simply believe that policymakers will be able to act for the best only if they understand what can and cannot be achieved.

In the opinion of this Group, **the way forward for evaluation** of the Framework Programme involves co-ordination of several elements:

- The natural focus on what is current or new should not remove the need to look back to what came before. It is as important to examine the outputs of the Fourth Framework Programme and the impacts of the Third Framework Programme as it is to obtain early feedback on the Fifth Framework Programme. At the same time we must recognise the changes in policy aims and instruments that have taken place during the lives of the three Framework Programmes.
- The present panel-based system for five-year assessments, without supporting studies and systematic data collection, is not equipped to properly assess economic and social impacts. **We propose a new basic model for five-year assessments.** Key elements of this model at the level of the whole Framework Programme include systematic use of assessments and monitoring reports from Specific Programmes, a review of changes in the environment of RTD, and an econometric study of meso and macro level effects. For Specific Programmes surveys and case studies should be supplemented by input of stakeholders view, analysis of sectoral trends and the use of indicators.
- Data collection should continue beyond the life of the project, and the standard contract needs to be changed to make this easier. To prevent evaluation from becoming an excessive burden, we should make sure that participants are asked for each piece of information only once.

- An external high-level expert group should provide guidance on methodological aspects of the evaluation scheme and help to ensure its quality and independence.
- Evaluation is in any case only one “intelligence” input to policy formulation. Along with other sources of information (for instance science and technology “foresight” initiatives, or studies of changing or emerging socio-economic needs), the results of evaluation studies should be debated in suitable public fora.
- Evaluation methodologies would become stronger if the Commission were to be more active in developing an evaluation culture in Europe.
- Proper evaluation needs a dedicated budget. We suggest that the Commission considers funding evaluation at 0.5% of its RTD efforts, as some countries already do.

## 1. The need for assessment

**Science and technology are driving forces in modern society. They contribute to many aspects of our public and personal lives, but often in complex and unpredictable ways that make it hard to assess their importance.**

Most of Europe's citizens, politicians and industrialists agree that science, research and technological development (RTD) are important for the economic and social development of our societies, and that much of the industrialised world's economic growth is due to RTD. Stakeholders argue about details such as whether advanced technologies can simultaneously provide both economic growth and jobs. Nevertheless there is a strong belief that competitiveness based on technology is a precondition of wealth, and that science and technology are essential if we are to cope with increasing environmental and ecological problems.

Public intervention and investment in RTD contribute to innovation and competitiveness, both in individual organisations and in society as a whole. Yet the relationship between RTD and the well-being of society is not simple, and experts agree widely that a linear input/output model is not an appropriate tool for analysing this relationship. In particular, we need to distinguish between long-term and short-term effects. We may expect that at least some of the long-term effects will spill over into society as a whole. Short-term effects such as increased profit are more likely to be confined to the participants in a particular RTD project.

Public spending needs not only to be controlled but also to be publicly accountable. Demonstrating the effectiveness of public spending on RTD helps to justify such spending and also provides a chance for policymakers to learn from experience. Public RTD programmes therefore need methods of evaluating both their immediate results and their longer-term consequences. The latter include direct and indirect effects, both social and economic, as well as unintended consequences.

Unfortunately the difficulty of attributing effects such as economic growth, job creation and social integration to particular public interventions seriously constrains our ability to assess the wider effects of RTD support. Another problem arises from the time lag between the RTD project itself and any socio-economic outcome. Furthermore, the results of innovation processes are largely unpredictable, and this severely limits any assessment of their potential impacts. This document explores the extent to which such difficulties can be addressed and managed.

This evaluation problem applies in particular to the European RTD Framework Programmes, since they — by definition — are supposed to fulfil functions supplementary to existing national RTD and innovation policy efforts, creating a “European additionality” of effects such as cross-border and inter-institutional networking of researchers and innovators.

Primary customers of any evaluation efforts of this type are the bodies politically responsible for such policies: the European Commission, the European Parliament and the Council of Ministers. Important secondary customers are the other central stakeholders of European RTD policies: research organisations, industrial companies and associations, consumer associations, and, last but not least, the general public.

This report focuses mainly on the assessment of RTD programmes rather than individual projects. It assumes that any such assessment will draw extensively on information about the impacts of individual projects, but that the impact of a complete RTD programme will be different from the sum of the impacts of its constituent projects.

In the coming pages we will discuss our basic understanding of any RTD programme assessment (Chapter 2); the challenges and limitations facing any attempt to measure the impacts, relevance and effectiveness of public initiatives (Chapter 3); the specific requirements of assessments of the European Framework Programmes (Chapter 4); and our opinions of the way in which future Framework Programmes should be assessed (Chapter 5).

## 2. Assessment: context, issues, definitions

**To gain insights into RTD impacts and their measurability, we must identify the impacts that can be expected, realise that they are related, and accept the complex nature of innovation.**

Europe's RTD policy is implemented via a large and complex instrument known as the Framework Programme. The Fourth Framework Programme (1994–1998) has a total budget of Euro 13,215 million. It covers all the major domains of science and technology, and by the end of 1997 it had financed more than 15,000 RTD projects to the value of Euro 8,200 million, while a further 10,000 projects had yet to reach the completion stage.

These projects involve 24,000 participants (including research groups, individual researchers, industrialists and universities) from all 15 Member States and several other countries. About one-third of the 25 Directorates-General of the Commission are involved in administering the Fourth Framework Programme. Any assessment of impacts therefore requires co-ordination at a large scale.

Several advisory bodies, both scientific and political, have been set up to help provide transparency and to ensure that the actual needs of the different stakeholders are taken into account. This complex system, managed mainly by the Commission Services, responds to changing conditions while remaining under the scrutiny of the Council of Ministers and the European Parliament. Bearing in mind the pervasive nature of research and the long time lag between the work itself and its socio-economic results, evaluating the impact of the Framework Programmes is an intricate issue.

Evaluation has been a legislative requirement for European RTD programmes since the early 1980s. In the years since then the Commission Services have acquired solid experience in evaluating research. For the first three Framework Programmes more than 70 programme evaluations and more than 40 supporting studies were carried out, involving more than 500 European experts in total. For the Fourth Framework Programme, the changing environment in science and technology and the increasing pressure for timely, independent evaluation has led to the current evaluation scheme, which has two components. First, each Framework Programme is subjected to *continuous monitoring* with the help of experts external to the Commission Services. Second, at multi-annual intervals external experts conduct a *five-year assessment* (Annex A).

The Commission's proposal for the Fifth Framework Programme sets out an even more ambitious evaluation scheme: *"The Commission shall continually and systematically monitor each year, with the help of independent qualified experts, the implementation of the Fifth Framework Programme and its Specific Programmes... It shall assess in particular whether the objectives, priorities and financial resources are still appropriate to the changing situation. Where appropriate, it shall submit proposals to adapt or supplement the Framework Programme and/or the Specific Programmes, taking into account the results of this assessment. (...) Before submitting its proposal for the Sixth Framework Programme the Commission shall have an external assessment conducted by independent highly qualified experts into the implementation and achievements of Community activities carried out during the five years preceding that assessment..."*

The funding criteria set out in the Commission's proposal ("European value added" and the subsidiarity principle; social objectives; economic development and scientific and technological prospects (see Annex B)) and, in particular, the shift towards broader socio-economic targets, consequently require a strengthening of the current mechanisms for assessment and data provision:

- The more the RTD objectives are expressed in terms of expected *socio-economic impacts*, the wider the scope of the assessment will have to be.
- As well as input data such as RTD expenditure and the characteristics of participants, more attention must be paid to gathering and assessing *output and impact data*. To reflect the full scope of potential socio-economic impacts, impact data need to be collected over a relatively long period of time after the completion of the project.
- The Fifth Framework Programme has been designed to be more *flexible* than previous Framework Programmes in its *implementation*. For this to work, decision-makers must be provided with timely information on the actual performance of the Framework Programme.

To learn about the potential and the limitations of RTD impact measurement and assessment, the European Commission convened our "Reflection Group" of experts charged with answering five questions:

- What are the *different kinds of impact* to be expected from public RTD programmes (taking into account both scientific and socio-economic objectives related to European policy)?
- What are the *limitations* facing our assessments of these impacts?
- What are the *implications* of these limitations on the evaluation of the impact of European RTD programmes, particularly from the viewpoint of the annual monitoring and five-year assessment exercises (under the particular requirements of the Fifth Framework Programme)? Taking into account the objectives and criteria of the Fifth Framework Programme and their potential implications for evaluation, what will be the most useful focus?
- Is it possible to *separate impacts* generated by the Framework Programme from those arising from other factors, including other research efforts? How is the *value added* by the Framework Programme best evaluated? What is the most appropriate *timescale* for such an evaluation?
- Taking into account the time and cost constraints, what are the *concrete first steps* in defining the required approaches? How can we *enhance* the present assessment scheme to best prepare for the next five-year assessment and other forthcoming evaluation exercises?

In compiling this report the Reflection Group decided that some clarification of terms was necessary. Within this document, therefore, we have developed the following definitions:

- **Monitoring** is a process by which information on the progress and direction of ongoing RTD actions is generated mainly for management purposes.
- **Evaluation** is a process by which the quality, implementation, target relevance and impacts of RTD programmes are investigated, interpreted and examined.
- **Assessment** is a synthesis of facts, which arise from the evaluation process, and judgements. Under this definition, assessment is a policymaking tool for the planning of new RTD programmes, not a political instrument.

### 3. Measuring impacts, relevance and effectiveness

**The key methodological challenges are rooted in the identification of different types of impacts, in their timing and in the fact that publicly-funded RTD is only one factor among many contributing to those impacts. This applies in particular to socio-economic impacts, which are difficult to trace to aggregate levels.**

This chapter deals with the generic issues facing any attempt to assess the effectiveness and impact of publicly-funded RTD initiatives. It examines the nature of the impacts to be measured, and outlines basic methodological requirements, with a focus on economic and social effects.

#### 3.1 The nature of impacts of publicly-funded RTD

As the targets of public RTD initiatives become broader and more heterogeneous, the more diverse and even contradictory impacts they are likely to produce. We can classify impacts in several ways:

*Tangible results versus knowledge and skills:* an RTD project may produce a tangible product, such as a machine, or an intangible outcome such as skills and knowledge. Skills may be thought of as knowledge embodied in human capital — a vital but often neglected result of RTD.

*Output versus impact:* to some extent it is possible to distinguish between outputs from RTD and impacts or effects which arise from the interaction between these outputs and the economy or society. Outputs can be purely scientific (such as publications), “intermediate” (such as patents and prototypes), or “final” (such as new or improved products, processes or services). Examples of impacts and effects are increased sales, improved competitiveness, and policies or regulations which improve the quality of life.

*Project versus programme:* individual projects may have been quite successful in terms of research results, technological solutions or economic results for the participants, while the programme as such may have led to unintended or even negative economic or social effects. Normally, however, the potential impacts of a programme exceed the sum of the effects of single projects.

*Participant versus non-participant:* impacts on participants in RTD programmes are inherently easier to study than those on non-participants, if only because the participants can be readily identified and normally have some obligation to co-operate with the evaluation.

*Core versus peripheral:* often, though not necessarily, related to the size of a participating company is the issue of whether the RTD is aimed at the core, or just the periphery, of the company's technological and business strategies.

*Continuity versus flexibility:* public RTD programmes normally run for some five years. Ideally, such a programme is clearly targeted and designed in a way that ensures straightforward and efficient implementation. The economic or technological conditions affecting the participants are less easy to define, and may change during the lifetime of the project. This may require flexibility in the way the programme is run, and will complicate the evaluation in terms of adherence to the programme's original aims.

*Short-term versus long-term:* impacts from RTD programmes appear over many different timescales. Short-term effects, for instance, may end abruptly as market conditions change. Other outputs may not be used for some years and then may become very important, perhaps because complementary technologies have been developed. Extrapolation is dangerous.

*Direct applicability versus unexpected applicability:* ideally, RTD project results can be applied directly in production processes or marketable products, thus contributing immediately to increased competitiveness. Even excellent research results and technological solutions, however, may become irrelevant to their planned markets if more competitive solutions are developed elsewhere. If this happens, the results still contribute to our "body of knowledge" and can sometimes lead to unexpected but successful developments that are completely different from the original aims of both the project and the programme.

*Economic versus social impact:* it is not always clear how to distinguish economic and social effects. Some economists argue that all social impacts can be expressed in monetary terms, for example through the "polluter pays" principle, but this involves judgements (such as the value of a human life or the cost of a reduced quality of life) that may not be widely shared. Other forms of impact such as employment are clearly in the economic domain, but have major social implications. In some areas, for example pharmaceutical research, there may be a strong positive correlation between wealth creation and quality of life. Other research, for instance on the harmful effects of smoking, would produce a strong negative correlation.

*Economic versus structural impact:* in Europe, structural effects take on an added importance since they manifest the creation of a "critical mass" in the Community, as well as being important for objectives such as cohesion. Again, though, their relationship with economic effects is complex. This could lead to an under-estimation of the expected economic impact, particularly because the benefits of networking are not easy to be quantified.

To a large extent, the *variety of impacts* and distinctions described above occurs because any RTD programme involves a *variety of actors* with differing interests and expectations: policymakers, programme managers, participating researchers and the programme’s clients. A well-designed evaluation has to take these different needs into account and to reconcile desires for information with constraints on resources and the availability of information. Table 1 shows *different classes of impact* (short or long-term; direct or indirect) in three *different domains*: the worlds of science, of the economy and society, and of policymaking respectively.

**Table 1: Impact dimensions of public RTD spending**

Main domains of impact of public RTD spending	Direct impacts		Indirect impacts	
	Short-term	Long-term	Short-term	Long-term
<b>Science</b> (“Wissenschaft”) Typical impacts:	scientific findings	knowledge	improved teaching	industrial spill-overs
<b>Economy and society</b> Typical impacts:	improved technology	improved technical know-how,	increased productivity	Improved competitiveness
<b>Policy</b> Typical impacts:	improved understanding	Problem-solving	increased problem awareness	increased general satisfaction

### 3.2 Methods and techniques for evaluation

According to a recent OECD report<sup>1</sup>, the proliferation and widening coverage of policy initiatives has led to evaluations that increasingly adopt a portfolio approach, and encouraged convergence between ex-post evaluation and continuous monitoring. This multi-faceted approach has become necessary because modern technology and innovation policies involve so many people — policymakers, programme managers, participants and evaluators — each needing different information.

We therefore need an integrated evaluation system that looks at the programme's rationale, economic impact, administrative conduct, customer satisfaction and other issues. There is growing awareness that there is no unique "best method" of assessment. Instead, we should combine several approaches. Examples include:

- qualitative and quantitative;
- ex-post, concurrent, and *ex-ante*;
- formative and summative.

Evaluation techniques which can be used concurrently fall into three categories according to their function<sup>2</sup>:

- *Framework methods* are used to place in context whatever is being measured and to make a judgement of value. Most framework methods involve an element of comparison (before and after, participants and non-participants, actual outcomes and expected outcomes), sometimes using logical frameworks. Also in this category are methods which measure the contribution that public funding makes to the overall results.
- *Data collection* is self-explanatory but non-trivial. It involves interviews, surveys, observation and ethnography, and the compilation of data and statistics from documents and other sources.
- *Data analysis methods* involve processing or structuring the data so as to test hypotheses or draw conclusions. These include case studies, econometric or other forms of formalised statistical analysis, and the construction of indicators and maps from statistics.

---

<sup>1</sup> OECD (ed): *Policy Evaluation in Innovation and Technology, Towards Best Practices*, Paris 1997

<sup>2</sup> Georghiou L., and Meyer-Krahmer F., Evaluation of Socio-Economic Effects of European Community RTD Programmes in the SPEAR Network, *Research Evaluation* 2(1): 5-15

### 3.3

### Focus on economic and social impacts

**We can only measure economic and social impacts if we know what we are looking for, and that depends on how we define these impacts.**

If we take *economic impact* to mean the contribution of a project to the competitive performance of industry, it might seem that the task of the evaluator is to measure how the project has increased profits for the participants. This could be measured as increased sales, increased market share or lower production costs, or a combination of these.

But increased profits cannot always be clearly attributed to specific projects, and even when a project has clearly helped to increase profits, the size of the contribution may be difficult to specify. Economic benefits may flow from market structuring (through the establishment of standards), from the development of skills and competencies which may be applied in other domains, through the provision of services and through all types of learning and networking effects. Just because its monetary benefits are hard to quantify does not mean that a project is not worthwhile.

An issue which needs to be addressed is the extent to which public funding has made a difference to a firm's R&D decisions. This 'additionality' question at its simplest explores whether the firm would have done the project in the absence of funding. The aim is of course to cause firms to increase their R&D spending, for example by undertaking projects with a higher risk but potentially higher payoff. Additionality can also be manifested by changing the way in which the project is done, for example making it larger, faster or more collaborative.

Non-participants can also benefit, either through deliberate or unintended knowledge transfer or simply from the commercial availability of the product or technology developed by the project. For collaborative RTD these more complex effects are likely to contribute a relatively high proportion of the economic benefits. The implication is that if the assessment is to be accurate, the assessor needs an in-depth understanding of the business processes and technology strategy of the firms taking part in the project.

Effects at the programme level may differ from those of individual projects; an obvious example is where a positive outcome for a project participant is offset by losses for non-participants. From the policymaker's viewpoint, however, it is generally an advantage to have a few strong players rather than many weak ones. Programmes may also affect the critical mass of researchers in a field in Europe or change the trajectory of that field. For these reasons economic impact may also be realised at a meso or macro-level.

*Social impacts* on employment, health, safety, environmental protection and public services are characterised by a large proportion of benefits that are not easily measured in money. As with economic impacts, what is good for the company may not be good for the community. An RTD programme may create new jobs in one company, but cause corresponding job losses elsewhere.

The benefits of a new road or rail link may outweigh the costs as long as we confine ourselves to economics, but the balance may not be so clear if we take into account wider effects such as the comfort or health of people living nearby. We can try to quantify these wider effects, for instance by adopting insurance companies' valuations life or limb, but there is little chance that everyone will agree on these valuations.

Although a few quantitative studies, based on a large public census, of the "value" of e.g. clean air or water were undertaken in recent years ("contingent evaluation"), the related methodology is still in an experimental status. So far, the best we can hope for is to expose some of the key trade-offs of RTD programmes.

Another complication is the fact that most social change is driven by factors other than RTD policy, for instance by immigration, taxation and education policies. For this reason it is normally impossible to attribute job creation and other specific social effects to a given RTD initiative, as a recent OECD document points out (see Box, next page).

In addition, RTD policies are characterised by an "impact gap": most socially-directed RTD does not address the problem directly but instead is used to develop the regulations, laws, standards or new practices which are necessary to bring about social change. If the "customers" of an evaluation or assessment exercise want information on employment impacts, they must be prepared to follow the suggestions of the OECD (see Box). The last sentence is especially telling.

### **Assessing the effect of technology policy on employment**

“The concern about high levels of unemployment in many countries has led to attempts to measure the employment impacts when evaluating technology programmes. Unfortunately, in many cases the political imperative of being able to point to job gains associated with some innovation and technology policy initiative has come at the expense of analytical rigor in the measurement of employment impacts. At the most general level, while there are clear and important links between technology development and employment creation, the links are seldom direct and not easily measurable. Hence, a requirement that a certain policy initiative aimed at encouraging innovation also leads to more jobs, while in principle a desirable aim, can in practice produce perverse results.”

“There are at least three pitfalls in such calculations. The first is a failure to distinguish between gross and net effects, i.e. to focus on the additionality implied by the policy initiative. Net effects are the difference between gross effects (the number of new jobs observed or forecast), minus dead-weight effects (the jobs that would have been created anyway in the absence of the programme), substitution effects (jobs that went to different people than would have been employed in the absence of the programme) and displacement or crowding-out effects (resulting when the policy initiative reduces activity and jobs elsewhere in the economy).”

“The second pitfall is the transition from direct to global job impacts. In the case of technology policies, job creation is rarely as important in the direct target firm, as it is in other parts of the economy. Such global job impacts occur through supplier effects (inter-industry sourcing of inputs), and through income multiplier effects (through the higher incomes that productivity-enhancing policies bring). Long-term supplier effects that occur when policy improves the knowledge base of the economy and its underlying growth rate, are also very important, but hard to capture.”

“A final hurdle relates to the translation from job creation to unemployment reduction. Ultimately, policy makers care about unemployment rates, and the level of unemployment is determined by the interaction of the demand and the supply of labour. Even policies whose net job impact is positive need to take into account their effect on the supply for labour, through their impact on activity rates.”

“The question of employment effects of technology policies will continue to preoccupy policymakers as long as job creation remains a problem. It is however important to be clear about what should be the objectives of different policies. In many technology programmes, employment objectives need not be directly identified. On the other hand, where such jobs impacts need to be identified, there is a need for more sophisticated approaches where information from surveys is validated by independent quantitative estimations and complemented by quantitative tools that capture economy-wide effects (i.e. input-output techniques, macroeconomic modelling or general equilibrium approaches).”

(Source: OECD 1998: *Technology, Productivity, and Job Creation: Best Policy Practices*, Chapter 5 “Policy Evaluation”, p. 135)

## 4. Focusing on the Framework Programmes: challenges and limitations

**An assessment of European RTD policy must take into account the specific objectives of the Framework Programmes: enhancing competitiveness and contributing to other Community policies while respecting subsidiarity and maximising “European value added” (for instance by contributing a small but critical part of the EU’s RTD spending, improving networking across Europe and encouraging European cohesion). The Fifth Framework Programme, considered as a “social contract” brings the additional need to assess the consequences of the Programme’s far-reaching socio-economic objectives.**

The previous chapters of this report have been generic: we have outlined the benefits and limitations of any impact assessment exercise that addresses the socio-economic effects of public RTD initiatives. In contrast, the unique nature of the Framework Programmes is the subject of this chapter. Framework Programmes are characterised by objectives that are closely targeted yet simultaneously defined in quite broad terms, and by firm principles guiding the selection of topics, programmes and projects.

### 4.1 Special features of the Framework Programmes

Community RTD initiatives are supposed to create “European value added”. In this context they must follow the subsidiarity principle, so as to select only those objectives which are most efficiently pursued at Community level. More precisely, projects should – beyond their quality in terms of scientific criteria, partnership, and project management either :

- contribute to create a “critical mass” of human and financial resources across all the Member States; or
- guarantee a significant contribution to the implementation of one or more Community policies; or
- address problems arising at Community level including social needs, or questions relating to standardisation or the development of the European area.

Moreover, Community RTD programmes and projects should contribute to the economy, science and technology in ways that will encourage the harmonious and sustainable development of the Community as a whole. This implies that projects will concentrate on areas in which:

- there is expansion and therefore good *prospects for growth*;

- community businesses can and must become more *competitive*;
- *scientific and technological progress* offers the prospect of medium- or long-term potential for dissemination and exploitation.

The First Framework Programme, launched in 1984, concentrated on industrial technologies, information technology, telecommunications and biotechnology. Each subsequent Framework Programme has been broader than its predecessor in its scope of technologies and research themes, with correspondingly higher expectations of its impact on the economy and society. As a consequence, the rationales underlying the various Specific Programmes under each Framework Programmes have become increasingly heterogeneous and even contradictory. This complicates attempts to evaluate the overall achievements of each Framework Programme<sup>3</sup>.

## 4.2 Widening the scope: the Fifth Framework Programme

In addition to the target dimensions applied already in earlier Framework Programmes, the Fifth Framework Programme will particularly emphasise social objectives that reflect the expectations and concerns of Europe's citizens. It has been claimed that the Fifth Framework Programme is a "social contract" which, much more than its predecessors, will aim explicitly to *create jobs, promote health and quality of life, and preserve the environment*.

Two other features of this "new approach" include an attempt to be concentrated and selective, with a limited number of programmes and "*key actions*", and the need for greater *flexibility* in programmes. The Fifth Framework Programme is divided into four "thematic" and three "horizontal" programmes.

The *thematic* programmes, which are concerned with RTD itself, are intended to be complemented by the *horizontal* programmes, which will promote co-operation, dissemination, and training and mobility of researchers. It is still difficult to say how radical a change in direction the Fifth Framework Programme represents, either in terms of scientific content or in organisation. The requirement to frame actions in terms of socio-economic impact and "European value added" is certainly a challenge to those responsible for defining programme objectives, work programmes, project selection criteria and assessment methods.

## 4.3 Consequences for impact assessments

The main difference between the Fourth and Fifth Framework Programmes is the Fifth's increased emphasis on wide-ranging socio-economic targets and a corresponding obligation to assess its achievements in these areas. Evaluators in previous Framework Programmes have had two decades to develop assessment structures to deal with at least

---

<sup>3</sup> See, for example, the *Five-year assessment of the European Community RTD framework programmes*, report of the independent expert panel chaired by Viscount E. Davignon 1997 (EUR 17644 EN).

some of the issues sketched in Chapter 3 of this report. The assessment of broader social impacts required under the Fifth Framework Programme presents new conceptual challenges.

As with most programmes, the Fifth Framework Programme also contains an element of trade-off between objectives. Although the objectives themselves will become increasingly verifiable, and the Specific Programmes and projects will be better focused than in the past, it is important to understand the logical connection between different levels of objectives. Some of the issues are:

- The broadening of objectives to include some which are specifically socio-economic implies an *increased number of stakeholders* and a *wider range of impacts*.
- The *participation of future new Member States* (especially in Central and Eastern Europe) will increase the variety of actors and research cultures involved in the Fifth Framework Programme and so make implementation and evaluation more complicated.
- The Fifth Framework Programme's *broad mixture of targets and expectations*, especially at the level of the Specific Programmes, is likely to cover the full spectrum of impact dimensions shown in Figure 1. As a consequence of the clustering of the Programme around "challenges and key actions", some of these heterogeneous impacts may also overlap.
- The potential reach of the various Specific Programmes is quite diverse. Although the Fifth Framework Programme's budget is small compared to the national RTD budgets of most Member States, EU programmes provide "*flexible money*" that can be fully invested in new projects. Most national funding, in contrast, is bound to particular institutions.
- Evaluating the *additionality* and added value of European RTD activities is very complex. For instance, we still lack effective ways to evaluate networking effects and those related to human resources.
- Existing and future micro-level evaluation studies are a potential input to more *macro-level analyses*. The latter are likely to become increasingly important in view of the expanded scope of the Fifth Framework Programme.
- Some socio-economic objectives of the Fifth Framework Programme, notably those concerning employment, may create *excessive expectations*. While job creation is a valid aim, employment is affected by many factors, policies and programmes outside the scope of the Fifth Framework Programme, and usually far more directly. Furthermore, the inter-relationships between these factors severely limit what can be achieved by an evaluation, however broadly cast.

A further general point is the need to respect the diversity of objectives in the Fifth Framework Programme. We need to accept the co-existence of multiple models of

innovation, their mixes differing from one sector to another. In turn, this will require *carefully-tailored “packages” of evaluation methods* to ensure that the impacts are adequately captured. This is not to say that comparability across areas is not possible, but rather that an attempt to compare on a single set of criteria will automatically bias the outcome.

Another problem arises from the time lag between the end of any RTD project and the appearance of measurable socio-economic effects, coupled with the fact that the rationale under which projects are supported and assessed changes somewhat from one Framework Programme to the next. Section 4.4 deals with this problem in more detail.

#### **4.4 The time problem: overlaps between Framework Programmes**

Successive Framework Programmes are conceived, administered and funded as separate entities, each implemented in the form of many thousands of RTD projects. Most individual projects have life-spans of two or three years, so projects started just before the legal cut-off date for a particular Framework Programme will finish well into the life of the next Framework Programme, and their socio-economic impacts will take even longer to emerge. As the first part of each new Framework Programme is largely spent on preparing new projects, it is probably appropriate to start each five-year assessment in the middle of the corresponding Framework Programme.

The assessment must, however, be handled with the utmost care; policy evolves more rapidly than RTD produces impact, and this should be taken into account when drawing conclusions about the effectiveness of the Framework Programmes. (A similar argument applies to the annual monitoring exercise, which takes account of results arising in the current year from activities funded three to five years previously.)

At any given time projects are active under more than one Framework Programme. In particular, two or even three Framework Programmes have to be assessed for short-term results and longer-term socio-economic impact at a time when a new Framework Programme is at the planning stage (Figure 2).

Overlap chart – another document

This feature, inherent in the nature of the Framework Programme, should be handled carefully to avoid misinterpretations or inadequate conclusions. The main potential risk lies in the execution of the five-year assessment exercise. According to its legal definition, the five-year assessment is conducted to support the preparation of the next Framework Programme. This implies that the criteria and frame of reference used to assess the results and impact of research depend on the objectives of the current Framework Programme, and tend to be strongly oriented towards future needs. As explained above, however, the most relevant information available at the time of the assessment comes from research projects initiated in the context of a programme that had somewhat different socio-economic and political objectives.

In conclusion, when using the outcome of a five-year assessment we must bear in mind that the multi-annual assessment system means that any study of impacts must work with at least three different “impact tracks” at the same time if it is to develop a balanced picture and allow learning from the past. For the forthcoming assessment under the Fifth Framework Programme, the long-term impacts of the Third Framework Programme will have to be studied at the same time as the short- and medium-term effects of the Fourth Framework Programme and the short-term results of the Fifth Framework Programme.

## 5. Assessing future Framework Programmes: the way forward

**The trend towards objectives framed in socio-economic terms requires a broader and more comprehensive assessment approach than is currently the case.**

Most reviews of RTD assessment practices soon discover that we are working at the *limits of known methods*. Despite the fact that we live in an increasingly knowledge-based society, the processes by which knowledge drives economic and social progress remain one of the most difficult areas to analyse.

The ambitious targets and complexity of the Fifth Framework Programme challenge the “assessability” of its impacts. The Reflection Group is convinced that the need for new assessment techniques will stimulate creativity and provide fresh and *innovative approaches*.

The Reflection Group therefore advises that we should maintain *an experimental approach*, using multiple methods as a matter of principle, while *continuing research* in this field, which is itself a manifestation of a socio-economic need. When setting objectives for evaluations we should recognise the limits of our assessment methods in terms of resources, timescale and methodologies.

### 5.1 Basic requirements: feasibility, intelligence and debate

Faced with the complexity of future European RTD policy assessment efforts, we can make five basic statements:

1. The ways in which Framework Programmes are assessed are *continually improving*. This positive process should be supported and strengthened, and we should avoid radical changes that would convulse the institutional, logistical or procedural basis of the current assessment system. Nevertheless, the system still reveals shortcomings in the way the impacts of RTD projects are analysed. The Commission can and should improve both the methodological and procedural aspects of its assessment practices.
2. All assessments and policy decisions are made with *incomplete information*; a fast but imperfect answer to a strategic question is often more helpful than a slow answer or none at all. Sound information is a prerequisite for any assessment, but increasingly detailed information seldom improves the quality of the decision.
3. Any overarching European approach to assessment should facilitate a “*multi-perspective*” view of the impacts of European RTD programmes.

It should take into account the differing, changing and even contradictory expectations of the various stakeholders, who include multinationals, SMEs, universities, contract research organisations, consumer associations and trade unions.

There is no single “objective” statement of the effects of policy initiatives; instead we have different interpretations of each programme’s targets and competing perceptions of its impacts. A learning-oriented, “*intelligent*” *assessment approach* would build on this diversity rather than trying to deny it. From this perspective, the main reason for assessing RTD programmes is to identify *changing requirements* in science, technology, society and the economy, and to adapt policy to these requirements.

4. If we are to use assessments to design future policy, we need fora in which to discuss these assessments and their consequences. The primary fora are likely to be the European Parliament, the European Commission and CREST, but we can envisage secondary fora at national or industry level. In either case, the resulting feedback is likely to be on two levels:
  - *Monitoring* at the level of Specific Programmes will help to give these programmes the flexibility required of them under the Fifth Framework Programme. Information of this kind will be used extensively by, for instance, the newly-created External Advisory Groups attached to the Key Actions.
  - The results of the *five-year assessments* — based on both monitoring and evaluation — will be used primarily as a resource for the design of future RTD programmes.
5. To successfully combine monitoring, evaluation and assessment exercises will require *tight co-ordination*, with particular efforts to avoid an inflation of evaluation reports.

## 5.2                    **Discussing a basic European assessment model**

Against this background, the Reflection Group has developed a **basic model** for any future assessment efforts in the context of European Framework Programmes (Table 3).

**Table 3 : Potential assessment steps for the Fifth Framework Programme**

**Table – in another document**

The model outlined illustrates the complex nature of RTD impact assessment. Clearly, there are many and different elements involved. Table 4 describes these elements in more detail and assigns priorities for practical purposes as regards the assessment of the fifth Framework Programme.

**Table 4A : Explanation to the suggested basic model :  
assessment of the Framework Programme impact**

<b>Activity</b>	<b>Priority 3)</b>	<b>Function</b>
<b>Framework Programme 5-year assessment</b>	Core	Panel convened to produce legally required report on Framework Programme on the basis of Commission's proposals for FP4 & FP5
Review of 5 year assessments of Specific Programmes	Core input	Ensuring that the findings of these assessments are systematically analysed, validated and synthesised by the FP-level panel
Review of Monitoring reports of Specific Programmes	Core input	Ensuring that the findings of the Monitoring Panels are systematically analysed, validated and synthesised by the FP-level panel
Review of changes in environment	Core input	Through consultation, review, (optional) studies and experience the panel should identify the main changes in the landscape affecting RTD in Europe, including policy, economy, industrial, social, legal & regulatory.
Econometric study of R&D in Europe	Recommended input	To provide evidence of the meso- and macro level impact of the FP as a complement to the bottom-up information collection of the other evaluation activities
Comparison with other (inter)national RTD policies	Optional input	With great care some aspects of administrative efficiency may be benchmarked in comparison with other programmes. Features such as priorities may also be compared.

3 *Core input* = indispensable element of every assessment; *recommended input* = important element of assessment; *optional input* = useful, but not indispensable, element of assessment.

**Table 4B : Explanation to the suggested basic model : assessment of the Specific Programmes' impact**

Activity	Priority 3)	Function
<b>Specific Programme 5-Year Assessment</b>	Core	Panel convened to produce legally required report on Specific Programme on the basis of Commission's proposals for FP4 & FP5
Results of monitoring efforts	Core input	Ensuring that the findings of the Monitoring Panels are systematically analysed, validated and synthesised by the Specific Programme-level panel
Questionnaire to ongoing projects (participants)	Core input	Questionnaire with generic questions to allow aggregation across programmes and specific questions which take account of the particular nature of the programme area and issues of concern to panel. Emphasis on strategy and management given incomplete nature of projects at this stage.
Questionnaire to participants in previous two Framework Programmes	Core input	Short questionnaire focussed entirely on economic and social effects for/from participants. For previous FP will normally be focus on outputs while returns from the FP before will provide concrete evidence of impacts.
Stakeholder/user/beneficiary views	Core input	Either through direct contact by panel, or more extensively through supporting study, aim is to identify views of relevance of programme for this group. Could be industrial, policy/regulatory, or social groups depending on programme aims.
Socio-economic impact studies on selected projects/participants (case or statistical studies)	Recommended input	Recognising complexity of tracing effects, this module focuses either on case studies or on statistical samples identified through data bases. In depth information should be collected either by interview or by elaborated questionnaires.
Sectoral studies of new technology, employment etc.	Recommended input	Using indicators to establish landscape in which programme is operating as aid to judge appropriateness of research strategy for that sector.
Scientometric studies of sector	Recommended input	Examining scientific performance of programme through citation performance. Also variety of other indicators including co-authorship, citation to patents, and mapping of research landscapes. Most relevant for scientifically-oriented programmes.
Effects on non-participants	Optional input	Seeks to identify positive and negative externalities arising from programme e.g. diffusion of technology or displacement of competitors. Difficult to implement. Alternative approach to examine performance of highly-rated but failed applicants as partial additionality test.

3 *Core input* = indispensable element of every assessment; *recommended input* = important element of assessment; *optional input* = useful, but not indispensable, element of assessment.

Taking the suggested basic model (Table 3) and its underlying components (Tables 4A and 4B) as a reference framework, the Reflection Group makes the following remarks as regards the implementation of the next five-year assessment:

(1) A basic requirement for the monitoring, evaluation and assessment of future Framework Programmes is the *systematic collection of data* at the level of individual projects, Specific Programmes and Framework Programme itself, and collected once the each project is complete. This data can only be collected through information at the project or participant level, and this raises several issues:

- The first is one of *efficiency*. Participants should not be required to deliver the same information to more than one study, whether these originate from inside or outside the Commission. The Commission has the primary responsibility for collecting such information, and could offer other stakeholders, especially the Member States, a data-processing service providing project information that is up-to-date and in a format suitable for their needs.
- Evaluators should have unrestricted access to all relevant documents held by the Commission which relates to projects. They should of course be held to account for maintenance and commercial confidentiality.
- Once efforts have been made to limit unnecessary inquiries, what remains should be adequately resourced and surveys carried out under *scientific conditions*, i.e. they must be piloted, confidential from programme line management etc.
- *Expert interviews* should be used to compensate for the limitations of surveys conducted by questionnaire.
- The greatest departure from standard practice is the need to recognise that many socio-economic effects take place over an extended period, mostly after the completion of the contract. A *tracking system* and in-built obligations or incentives should be used to ensure that data can still be collected from participants during this extended period. The *standard contract* will need to be amended to guarantee the provision of appropriate data in the period following completion.

(2) The structure and the quality of the *data* will determine whether it can be *aggregated* for the purpose of evaluation at the levels of the Specific Programmes and of the Framework Programme itself.

(3) The *reporting frequency* for monitoring and assessment could be reconsidered. External monitoring reports should probably be issued every two years, instead of every year as at present. This would avoid bothering the beneficiaries too often, and would focus attention on larger-scale changes in the development of the project and its parent programme. It would also help to save the Commission's resources.

“Evaluation fatigue” can be a problem with project participants who have to respond to many different questionnaires, especially when these ask repeatedly for the same data. Evaluation fatigue could be avoided by co-ordinating the gathering of data for

monitoring and evaluation studies with the information needs of other RTD policymaking bodies. For example, “Impact Studies” on the role played by European RTD programmes in national innovation systems could use existing Commission data. These studies could then feed their own results back into the Framework Programme assessment system.

(4) The *panel system* used by the Commission to conduct its assessments successfully addresses the need for demonstrated independence. However, since panels do not have a monopoly of wisdom, supporting studies should be conducted and their results made generally available. The assessment panels will in any case be required to make judgements on the broader strategic issues raised by these findings. A separate, *methodologically-oriented, panel* could also be set up to oversee the process and guarantee its quality.

To increase the amount and quality of information available to the monitoring panels, and to enable them to cope with the new challenges posed by the Fifth Framework Programme, the panels could be given the opportunity — in addition to participating as observers in the project selection process — to attend conferences, workshops and other events organised by the programmes. This would allow panel member to meet project participants and other actors, and could be an effective way to further improve confidence in the monitoring reports.

(5) Any “intelligent” assessment of RTD impacts by panels needs supporting information which to a great extent can be achieved from *independent studies*. Information may be needed on:

- future scientific and *technological developments*, sketched by foresight studies;
- changing *socio-economic needs* of particular industrial sectors or societal groups, described by in-depth studies in economics and social sciences;
- *specific impacts*, intended or unintended, of funded RTD activities, analysed by thematic evaluation studies or by technology assessment efforts; and
- the *appropriateness* of programme *designs* and methods of *implementation*, investigated by studies of policy or management.

Supporting studies may be necessary as an input to the assessment of Specific Programmes as well as of the Framework Programme in general. In each specific case it is crucially important to identify the best evaluation system, not so much by “inventing” new methods as by carefully exploring existing approaches and choosing the most appropriate.

An example is the bibliometric or technometric “mapping” of RTD directly related to Key Actions. This is a powerful tool for visualising the cognitive landscape of an RTD field and its surrounding environment. Maps made over a series of years reveal trends

and changes in structure, and maps can be extrapolated to act as a “foresight” system for near-term RTD developments. Related quantitative (bibliometric) and qualitative (peer surveys and user surveys) methods can also be used to map the positions and strengths of the main actors. Changes in these maps over time can then be used as evidence of the impact of RTD programmes.

(6) A demanding assessment approach as discussed in this report needs a *dedicated budget*. Programme managers can only organise monitoring, evaluation and assessment procedures appropriately once certain amounts have been explicitly earmarked for assessment efforts. In other contexts a budget of 0.5% of the RTD funds allocated has shown itself to be sufficient, so there is clearly still plenty of scope for extending the monitoring, evaluation and assessment budget of the Framework Programme. Such a direct allocation of financial resources is also in line with the general recommendation formulated by the Commission in its communication on the diffusion of evaluation best practices (SEC (96) 659/5, p.14), which advocates that 0.5% of total expenses should be earmarked for programme evaluation.

### **5.3 RTD assessment in a broader European context**

Assessments of the achievements of European RTD initiatives should be complemented (where this is meaningful, such as at the level of Key Actions) by comparisons with the results of other RTD policy initiatives. These may be *international* (such as EUREKA and COST), *national* or *regional*. Different programmes have different objectives, so “comparison” should not be interpreted as “benchmarking”, and the temptation to rank the achievements of different programmes should be resisted. Nevertheless, a benchmarking approach might be useful when comparing the management performance of the Framework Programmes with that of other initiatives.

To facilitate assessment in the way we have outlined in this report, the underlying “*evaluation culture*” in Europe needs a process of continuous improvement. The credibility of RTD policy evaluation and confidence in its outcomes and use could be strengthened if assessment results were generally accessible to the interested public and could be debated regularly in suitable fora. This could at the same time raise public awareness of EU RTD policy, and ultimately increase the social and economic benefits of this policy to European society.

## **Annexes**

- A. Overview of the EU RTD Programme Monitoring and Evaluation System
- B. Criteria for Community RTD activities
- C. Legal basis for Fifth Framework Programme and related evaluation

