

The Nature and Scope of RTD Impact Measurement

A discussion paper for the International Workshop on “Measurement of RTD Results/Impact”, Brussels, 28-29 May 1998

Dr. Laurent Bach¹ and Prof. Luke Georghiou²

1. Introduction : Purpose, Scope and Criteria

To plan any evaluation the necessary starting condition is clarity in terms of the purpose of the evaluation, of the scope of the activity to be evaluated and of the criteria to be employed. In addressing the subject matter of this workshop a certain focus emerges. While it is clear that they are not independent from one another the interest here is upon socio-economic effects or impacts of RTD, not upon the scientific quality of research outputs, nor upon the management of the programmes which seek to deliver these effects. The history of evaluation of the Framework Programmes has shown that the latter two dimensions have posed significantly less of a challenge for evaluators. Scope is also implicit in the concern with RTD programmes. There are many other dimensions of policy which impinge upon technological innovation and diffusion deliberately or otherwise. While these provide an important context for the exploitation of RTD results, an issue discussed below, in themselves they have traditionally demanded different approaches to their evaluation which will not be addressed here.

The criteria for the evaluation of the Framework Programme have mirrored the objectives of the Programme itself and hence in terms of impact have concerned the effects upon the competitiveness of European industry and contribution to other Community policies. These factors of course still apply for the Fifth Framework programme but there are some changes which potentially have implications for evaluation. In common with a more general international trend the new Programme is framed explicitly within the context of socio-economic objectives. The term socio-economic is chosen deliberately to emphasise the presence of the social dimension alongside the economic. These include employment, quality of life and health and the environment and are manifested in unifying themes such as the ageing population. Not only do these objectives imply a multi-disciplinary approach to their implementation, but also an evaluation system which can distinguish and measure benefits which may not be monetary.

In considering the scope, a key issue to be addressed is that of the level of aggregation. RTD programmes may be conceived in terms of macro or meso level objectives but

¹ BETA (Bureau d'Economie Théorique et Appliquée), L. Pasteur University of Strasbourg, 38 bd d'Anvers, 67000 Strasbourg, France.

² Director, PREST (Policy Research in Engineering, Science and Technology), University of Manchester, Oxford Rd, Manchester M13 9PL, United Kingdom

their implementation is firmly rooted at the micro-level. Evaluation has remained at this level because of the argument that the Framework Programme has accounted for only a small proportion of spending upon on RTD within Europe and hence that its macro-impact is very likely to be lost in the noise of effects arising from other inputs, for RTD and beyond. This is perhaps an over-pessimistic attitude. If it can be established (as is likely) that Framework funding leads to an increased level of RTD spending within Europe (or even a more effective use of constant resources), then some estimate may be made of the macro-economic effects. A recent calculation made in Japan, based upon a complex model, has estimated that the present plans to double government investment in RTD will add 1% to the growth rate of the Japanese economy in the next century. Such an assurance, even with the necessary caveats, would provide a welcome complement to the present “bottom-up” approaches. Construction of such a model would need to be built upon a clear understanding of the effects of Framework support upon the behaviour of firms, as well as upon their RTD investment. One key question here is to move from a purely quantitative relation between money invested in RTD in general and GNP or other macro-economic indicator towards a macro-economic analysis of the impact of the support to RTD provided through one specific institutional arrangement that is called the FP.

In any event, meso and macro level effects are unlikely to be the arithmetic sum of the effects of projects. On the negative side issues such as displacement effects upon other European firms as a result of the improved performance need to be considered. On the other hand the net effect of a programme may be to create a critical mass of research in an area which leads to some form of growth cycle, for example by creating a labour market for particular specialised skills emerging from the new technology or by attracting inward investment.

Finally, in the way of introductory remarks, the question of the purpose of evaluation needs to be addressed. A complex policy instrument such as the Framework Programme has multiple stakeholders, and while all are concerned that desired impacts should be achieved, their information needs are different. Considering these stakeholder groups in turn:

Participants - it could be held to be self-evident that participants in the programme are aware of the impacts arising from their RTD activities but this is not necessarily the case. Participants are a varied group and for industrial participants, it may be necessary to consider separately the motivations of researchers from those directly involved in markets and corporate decision-making. Even among large firms there is an extensive debate in the literature as to ways in which performance metrics may be applied to RTD, with no dominant approach emerging to date. The methods which are used are recognised to be at their most effective for measuring the benefits of applied problem-solving research and least effective for the types of research pursued in collaborative programmes. In any event, firms are less likely to be aware of impacts upon other actors. The benefits to firms of adequate impact evaluation are twofold: the opportunity to learn and manage their resources in amore effective way by re-inforcing success, and obtaining evidence which motivates senior management to maintain or increase RTD investment.

Programme managers - programme managers' interest in the impacts achieved at project and programme level mirrors that of participants. Understanding the linkages between actions and effects improves the selection and management of future projects, while broader evidence of the impacts sustains the case for public resources to continue to be invested in their particular domain.

Policy-makers - Policy makers may be considered to include both those responsible for implementation at a higher level and those whose function is to represent the broader interests of citizens. For policy makers the arguments of learning and justification apply at a higher level of aggregation but to these are added the question of public accountability, that is the need to demonstrate to the European taxpayer that value-for-money has been obtained. With the progressive raising and broadening of the level at which objectives have been stated (through competitiveness to social goals) it is likely that more explicit evidence will be needed to demonstrate the relative effectiveness of RTD support against other measures in the pursuit of these goals. Implicitly, the broadening of the objectives statement makes clear the role as stakeholders of bodies not previously directly concerned with RTD (for example policy makers and pressure groups in the domain of care of old people).

2. Impact and Innovation

Before considering either specific impacts or the methods by which they may be measured it is critically important to emphasize that methods are rooted in paradigmatic views about innovation and the economy. It may be argued that in some sense impacts are "created" by the methods and the corresponding theoretical background. To elaborate, consider first the linear model of innovation. In essence this model assumes the existence of a series of steps, sequentially linked in which the output of each becomes the input for the next. At its crudest this model is represented by a research phase in which scientific knowledge is produced, following which there is a transfer to a development environment and then a transfer of a prototype or its equivalent to a context of commercialisation. For evaluation, the implication is to consider that at each stage, there are identifiable and measurable outputs which exhaustively represent the contribution of this stage to the innovation process. Another implication would be to identify the final stage where monetary or social benefits are incurred and then, according to the methodological choice, either to associate the full range of benefits with the original research input, or else to work back through the sequence calculating a co-efficient governing the input-output relations.

In an analogous fashion, a classic approach to RTD distinguishes between the results of basic or fundamental research and results from downstream RTD activities; the first are seen as public goods and the latter as private goods, with each treated according to separate theoretical frameworks and methodologies in dimensions such as appropriability.

On the other hand more recent approaches to innovation, grounded in interactive, evolutionary and systematic contexts, identify the linear model as an extreme special case which by neglecting feedbacks leads to a systematic underestimation of the benefits of RTD. Features such as learning phenomena, structural changes induced by knowledge generation, establishment of standards and the creation of networks all

enter the frame of analysis. Differences between technologies and markets become crucial in understanding impacts. Such differences include appropriability, often founded in the degree to which knowledge is tacit or codified in a particular sector and leading to different intellectual property protection strategies. With a more complex view of appropriability the distinction between private and public goods becomes less significant. There are several implications for evaluation, including the need to give appropriate weight to the contribution of intangible outputs to the effects, and the need to understand the relationship and interactions between outputs at all stages of the innovation process. Knowledge generated through RTD is just as likely to be applied to a problem arising from a technological development at the market stage as it is to the next step in a putative sequence.

A focus on a particular measure creates an implicit weighting among the variables. Hence, a concern with monetary evaluation highlights those benefits captured by price mechanisms at the expense of externalities and effects on policy or social factors.

Recent discussion of *ex ante* valuation of RTD in an industrial context has focused upon approaches based upon option-pricing - that is to say that investment in a research project may be thought of as analogous to the purchase of an option to buy future stock. In other words, by performing the project the opportunity to progress in a particular direction is created. Without addressing the full complexities of the method, for the purposes of this paper it is worth noting that an options perspective generally leads to a higher valuation of research than a straightforward net present value calculation. Furthermore, this way of thinking can be extended to higher levels of aggregation. Choices made in a particular programme may influence the direction of a technological trajectory (consider the emergence of the GSM standard in mobile telecommunications) or indeed the capabilities in a field at national or European level. However, similar problems to those discussed above are encountered in the move from micro to macro.

3. Impacts of RTD

3.1 Dimensions of Impact

In considering impacts there are several dichotomies which may be applied:

Artefact vs. knowledge and skills - concerning whether what is produced is tangible or embodied (product, process, service) or intangible in terms of knowledge and skills (which may be thought of as knowledge embodied in human capital - a vital but sometimes neglected result of R&D which can be manifested at the individual, organisational or inter-organisational level and may be tacit or codified);

Output vs. impact - to some degree it is possible to distinguish between *outputs* from RTD, (ranging from scientific outputs such as publications, through “intermediate” outputs such as patents and prototypes, to “final” outputs such as new or improved products, processes or services) and *impacts* or effects which arise from the interaction between the outputs and the economy or society (for example sales, improved competitive position, or policies/regulations which improve quality of life).

Short term vs. long term - there are many different time-profiles over which impacts are manifested including short-term effects which terminate abruptly as market conditions change and outputs which are not used for some years and then become very important, perhaps because complementary technologies have been developed. A key point is that extrapolation is dangerous.

Intended vs. unintended - excessive focus on project or programme goals as the basis for evaluation could lead to important unintended effects being missed - this effect can be exemplified in medical research when a drug developed for one condition turns out to be important in the treatment of another. Unintended effects also include most of those in the negative category.

Participant vs. non-participant - impacts upon participants in RTD programmes are inherently easier to study than those upon non-participants, if only because the first group are readily identified and normally have some obligation to co-operate with the evaluation. Among participants the impacts are generally dependent upon their role in a project. Hence, a component supplier would be seeking a new product to sell, a user would seek to reduce process costs or improve quality through application of that product, and a knowledge supplier would be seeking to improve its knowledge base and reputation in order to win further research contracts. The relative weight of different outputs and impacts would be different in each case.

Non-participants could be affected in several ways:

- through reception of outputs, effectively the consumer surplus embodied through the purchase of outputs or else the diffusion of knowledge to non-participants through deliberate disclosure (publication, patents and standards), demonstration or imitation effects or leakage through informal networks, mobility of personnel etc.;
- through *competitive relationships* with the participants, in terms of loss of market share to a successful participant (described above as ‘displacement’), or more positively through a competitive response to the improvement made by the participant, inducing a general raising of the level of the sector; and
- through effects which *induce broader structural or regulatory changes* affecting the sector; and
- through other network relationships with non-participants.

Core vs. peripheral - often but not necessarily related to the size of a participant is the issue of whether the effect is upon the core or the periphery of the technological and business strategies of the participant.

Economic vs. social impact - it is not always clear in which category a particular effect may fall. Economists may argue that ultimately social impacts may be captured in a monetary way, for example through the polluter pays principle. However, such transformations often involve arbitrary translations (the value of human life or of reduced quality of life) which are not supported by a social or political consensus. Other forms of impact such as employment are in the economic domain but have major social implications. In some areas, for example pharmaceutical research, there may be a strong correlation between wealth creation and quality of life, while other research (say on the harmful effects of smoking) would produce a strongly negative relation.

Economic vs. structural impact - particularly in the European context, 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 the relation with economic effects is complex. At one extreme the formation of a new network may only serve as a device to consume public funds, while towards the other end of the scale, networks are becoming of increasing economic importance, with external linkages often being a pre-condition for innovation in many areas. As well as the 'temporary' network required to execute the RTD project, a linkage of this kind may form all or part of a network which performs further functions in the technological or market domains. Since the cost and risk of forming new linkages is high (evidenced by the very high level of repeat linkages in successive programmes) there is clearly added-value in stimulating their formation.

3.2 Taxonomies of impacts

As well as the dimensions discussed above, almost every evaluation, or approach to evaluation generates its own list of impacts. Table 1 below shows a list derived from the COMEVAL Toolkit, selected not because it is better than the others but because it is one of the longer lists available.

Outputs		Impacts/effects	
Intermediate outputs	prototypes technological sub-systems demonstrations models/simulators integration of technologies tools/techniques/methods intellectual property decisions on further RTD	Competitiveness	sales market share open up markets create new markets lower costs faster time to market licence income
Products	new products	Employment	jobs created jobs in regions of high unemployment

Processes	improved products new processes		jobs secured jobs lost
Services	improved processes new services improved services processes for delivering new services	Organisation	formation of new firm joint venture to exploit results new technological networks/contacts new market networks/contacts improved capacity to absorb knowledge core competence improvement further RTD change in strategy reorganisation of firm to exploit results increased profile
Standards	de facto standard de jure standard reference conformance memoranda of understanding common functional specification code of practice identified need for regulatory change	Quality of life	healthcare safety social development & services improved border protection & policing support for cultural heritage
Knowledge and skills	management & organisation technical	Control & care of the environment	reduced pollution improved information on pollution & hazards reduced raw material use reduced energy consumption positive impact upon global climate decrease in pollutants
Dissemination	training activities workshops/seminars/ conferences technology transfer activities knowledge & skills transfer publication/documentation	Cohesion	employment in LFRs infrastructure of LFRs participation of LFRs further RTD in LFRs regulation and policy in LFRs
		Development of infrastructure	transport telecommunications urban development rural development
		Production & rational use of energy	energy savings renewable sources nuclear safety assurance of future supply distribution of energy
		Industrial development	development of internal market development of SME sector development of large organisations support for trade
		Regulation & policy	EU regulations or policy national regulations or policy world-wide regulations or policy co-ordination between national & Community RTD programmes

It should also be remembered that the precise classification of an output is context sensitive - a firm may sell prototypes as its main business. many of the sub-categories shown themselves break down into further multiple sub-categories. Furthermore this particular taxonomy was designed to be operational at the project level, though almost all of the effects shown are also manifested at the level of the organisation or above.

4. Barriers to Assessment

A series of key issues complicate the evaluation of the impacts of an RTD programme. The first is that of separability of the evaluated programme and the evaluated effect, also known as the problem of attribution. In essence the problem is simple - an effect may arise from multiple causes. These may include other R&D projects within the programme or outside it and also include inputs to innovation other than R&D (for example external technological sourcing). A relatively common situation among large firms is that of the “project fallacy” whereby policymakers see the contract which they fund as an entity which produces discrete effects when as far as the funded firm is

concerned it is either a part of a larger programme of work (either longer or broader) or else requires substantial follow up work and combination with knowledge from other sources before an economic effect is realised. The key distinction is between a contract and a project. A contract offers deliverables in return for a grant and has a fixed life. In collaborative RTD these are joint deliverables from a team of partners. This contract, from the point of view of a participant, is a means by which to further the goals of the organisation, which for a firm are normally its commercial goals. Normally, both the grant and the benefits of collaboration help the firm in this way. However, the effects (for example increased sales) arise from the broader/longer work programme mentioned above. Successive attempts by evaluators to ask firms to calculate a rate of return for a particular project typically meet with the response that this is not a meaningful question - the rate of return is one on the entire innovation cost. An analogous situation is true for social effects. There are two implications for the evaluator:

- the contracted RTD has to be situated within an understanding of its contribution to the technology and business strategy of the exploiting organisation;
- effects are best measured in terms of a null hypothesis, i.e. what would have happened in the absence of the project.

Following on closely from this is the issue of the additionality of the policy intervention, in this case the contract to perform collaborative RTD. This is a question which can be addressed in more than one way:

- would the firm have done the project without the contract?
- would the firm have done the project differently without the contract (faster, larger, collaboratively)?
- has the existence of the project/programme influenced the technology strategy of the firm (e.g. taken it into new areas)?
- has the firms R&D expenditure increased by an amount equivalent to the value of the project (or greater, if public and private funding are complementary)?

All of the above questions are difficult to answer definitively. It is not obvious whether projects with high additionality will produce greater or smaller impacts. On the one hand one would expect firms to cover their highest priority projects with their own resources and hence put forward marginal projects for funding. This rationalistic analysis is not confirmed by the evidence from evaluations - there are ample examples of projects with both high additionality and high subsequent impact. One explanation could be that public funding motivates firms to undertake projects with a higher risk but potential higher pay-off. Another aspect of this question is the possible appearance of non expected technological and/or commercial opportunity during or after the project, which may transform marginal project to higher priority projects. The question about whether the RTD would have been done differently raises a further point, that of the *persistence* of effects. If behavioural changes are stimulated by the intervention are these maintained beyond the period of the intervention. Success in the Framework Programme's structural objectives are dependent upon achieving a lasting shift and all other effects are increased to the extent that the firm has learned from its experience. The issues raised here are an essential component of an evaluation.

The other main barriers to evaluation of impacts have already been raised above. Limits to the appropriability of knowledge generated, and the consequences of realised effects upon non-participants mean that externalities have to be included in a comprehensive assessment of impacts. Timing creates a practical problem. While some structural effects and spin-offs may have been realised by the end of the project, many others will take time to be manifested. While it may be argued that the Framework Programme does not support basic research, even nearer to the market RTD can take years to be “productised”. As remarked above, expectations at the end of the project are not always accurate and hence a follow-up mechanism is necessary. This has practical implications for evaluation. A suitable period might be three years after the end of the project. It would be prudent to include in the original contract some provision for post-project reporting. Furthermore, over that period the turbulence of corporate re-organisation may make original contacts difficult to trace. If commercialisation is in process the original research managers may no longer be the appropriate contacts within the organisation. The implication is that a significant effort is needed.

5. Measurement and Interpretation

In view of the difficulties discussed above, how can an evaluation proceed to the measurement of impacts ? It should already be clear that a checklist approach is unsatisfactory. Without a clear mode of the dynamics of the participation or project under consideration a highly misleading picture could be constructed. Simple aggregations of outputs mean little without some contextualisation - comparison with a suitable benchmark figure, adjusted for the type of firm, the region, the country, the scientific area, the industrial sector....and so on.

Movement from an impact to an indicator useable for evaluation purposes also depends upon the scalability of the selected item. While attributed effects on sales and other such monetary indicators can be aggregated or compared, effects on networks are much more difficult to assess. There have been some useful attempts to map the growth of networks but these do not go onto provide a quantitative measure of the utility of those networks to their members. Skills and competences in general are difficult to measure except in terms of their aggregate effect upon the performance of the firm. Case study evidence yields statements such as “we did not achieve our technical goals but the project gave us a competence in that software methodology which has enabled us to win new contracts”. The credit to the programme is in the range of zero (skills was not an objective of the project) to the total value of the new contracts (which were also dependent upon other dimensions of the firm’s performance).

At a practical level, the source of information has to be considered. The dominant approach is that of reporting by the participants, either through monitoring procedures or in response to evaluation questionnaires. It would be a brave evaluator that admitted that a questionnaire could be relied upon to capture the complexity of the innovation process as revealed by an in-depth interview (or series of interviews) leading to a case-study. How to manage the trade-off between cost and quality is a central design issue for evaluation. Whichever approach is chosen to obtain information from participants,

it is context sensitive. Without confidentiality from those who might make future decisions affecting the participants it can be assumed that some bias is introduced in the response. This is not necessarily a predictable bias - the type of state-aid which requires successful projects to repay all or part of their subsidy is likely to lead to a downplaying of benefits, while a system in which future funding goes to those with a successful evaluation rating will stimulate exaggeration. As far as possible evaluations need to seek means of verification to improve the reliability of the data, for example by comparison with statistics collected for other purposes, or by accumulating data throughout the life of the project and beyond to avoid the temptation to hindsight. All of this has to be in the context of minimising the burden upon a population which may already be experiencing "evaluation fatigue".

Also from a practical point of view, the question of the clarity of the evaluation work and the evaluation process is crucial, in the sense that the results should be properly understood, compared and interpreted by the different stakeholders mentioned in the introduction. This means for instance that a relevant balance should be found between different evaluation methods and practices, allowing on the one hand recurrent evaluations to be carried out, with repeatable methodologies and systematic collection of data (for instance mainly short term S&T and market related indicators, with the largest coverage in terms of number of participants and projects), and on the other hand evaluations encompassing more broad views and combining qualitative and quantitative, structural and economic, mid-term and long term, etc, perspectives. This also requires correct and so-to-speak modest awareness of the scope of each evaluation work and of the significance of the results obtained (for instance by providing conservative estimation for quantitative evaluation).

6. Conclusions

To conclude, some of the points made above are considered in terms of their relevance to the evaluation of the Fifth Framework programme. A first point is that the broad scope of the objectives of the Programme, while providing a valuable orientation for the research itself, presents a challenge in terms of operationalisation for the evaluator. Criteria for the Programme include selection criteria, subjective aspects such in the realm of quality of life and effects which are difficult to integrate in an analytical frame. As with most programmes there is also an element of trade-off between objectives. While objectives will no doubt become increasingly verifiable as the level of detail increases through specific programmes and projects, it is important to understand the logical connection between each level.

The second point is the need to respect the diversity of objectives in the Programme and hence to accept that multiple models of innovation will co-exist, the mix depending upon the sectors in question. In turn this will require careful tailoring of a "package" 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.

All of the methods employed demand information at the project or participant level, raising several issues. The first is one of efficiency - ensuring that participants are not required to deliver the same information in slightly different formats to studies initiated

at different levels of responsibility or location, inside or outside the Commission. On the other hand, once efforts have been made to limit unnecessary inquiries, what remains should be adequately resourced and surveys carried out under scientific conditions, (piloted, confidential from line management etc.). The limitations of surveys require that a substantial sample of expert interviews are also used. Perhaps the greatest departure from standard practice is the need to recognise that effects take place over an extended period most of which is after the completion of the contract. This implies a tracking system and an in-built obligation/incentive for participants to continue reporting. Almost certainly the standard contract will need to be amended.

The Panel system employed by the Commission to conduct its evaluations successfully addresses the need for demonstrated independence but it does not sit easily with the need for complex information about impacts. The responsibility rests with the evaluation service and experts they commission to provide such information for panels to comment upon. Since panels do not have a monopoly of wisdom in such interpretation the studies should also be made generally available. Perhaps a separate methodologically-oriented panel could be employed to oversee the process and offer it a guarantee of independence in the Chabalian sense. The panels will in any case be required to make judgements on the broader strategic issues raised by these findings.

As a final word, any review of evaluation of RTD soon discovers that we are working at the limits of known methods and approaches. Despite living increasingly in a knowledge-based society, the processes by which knowledge drives economic and social impacts remains one of the most difficult areas to understand. In these circumstances the Commission would be well advised both to maintain an experimental approach, using multiple methods as a matter of principle, and also to recognise the need for continuing research in this field which is itself a manifestation of a socio-economic need.

