EUROPEAN
COMMISSION

Community research

What the Evaluation Record tells us about the Framework Programme Performance

EUR 21943

Interested in European research?

RTD info is our quarterly magazine keeping you in touch with main developments (results, programmes, events, etc.). It is available in English, French and German. A free sample copy or free subscription can be obtained from:

European Commission
Directorate-General for Research
Information and Communication Unit
B-1049 Brussels
Fax (32-2) 29-58220
E-mail: research@cec.eu.int
Internet: http://europa.eu.int/comm/research/rtdinfo/index_en.html

EUROPEAN COMMISSION

Directorate-General for Research
Directorate A — Coordination of Community activities
Unit A.4 — Planning, programming, evaluation

E-mail: rtd-evaluation@cec.eu.int

European Commission
B-1049 Brussels

EUROPEAN COMMISSION

What the Evaluation Record tells us about the Framework Programme Performance

Erik Arnold

2005 Directorate-General for Research EUR 21943

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number:
00 800 6 7 8 9 10 11

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://europa.eu.int).

Cataloguing data can be found at the end of this publication.

Luxembourg: Office for Official Publications of the European Communities, 2005

ISBN 92-79-00385-6

Printed in Belgium

PRINTED ON WHITE CHLORINE-FREE PAPER

This paper reviews the entire record of mid-term and ex-post evaluation of the Framework Programmes in the period 1999 to mid-2004 and provides a synthesis of the main findings and messages. It is part of the "knowledge-base" underpinning the Five-Year Assessment of the European Union Research Framework Programmes 1999-2003, which was carried out by a high level independent expert panel in the second semester of 2004.

This publicly available collection of nearly 150 documents includes 22 Community assessments or evaluations, 7 evaluation policy and methodology documents and 12 national impact assessments or evaluations. It also contains 69 policy documents and reviews and reference documents such as previous monitoring or Five-Year Assessments, Annual reports on research activities (art. 173), indicators and the Framework Programmes' legal base.

These documents are available on
http://forum.europa.eu.int/Public/irc/rtd/fiveyearasskb/library.

WHAT THE EVALUATION RECORD TELLS US ABOUT THE FRAMEWORK PROGRAMME PERFORMANCE

Technopolis
Erik Arnold

January 2005

Contents

What the Evaluation Record tells us about Framework Programme Performance

Summary

This paper reviews the record of mid-term and ex post evaluation of the Framework Programmes in the period 1999 to mid-2004 and distils key messages from it. The analysis served as an input to the Five-Year Assessment 1999-2003 of the EU RTD Framework Programme (FP). The Five-Year Assessments form part of the regular programming cycle for EU R&D programmes.

The Framework Programme provides a broad and permissive context for RTD programmes. Its high-level goals are to strengthen the research capabilities underpinning European industry and to improve citizens' quality of life, while its low-level goals are largely to do R&D in its constituent Specific Programmes.

The available evaluation evidence suggests that the Framework broadly funds good quality work, in which universities and research institutes play a majority and increasing role. Framework participation is led by a 'core' of major beneficiaries, who sit at the heart of multiple European RTD networks. There is scope for greater industrial participation, which could be desirable in order to reach the Barcelona goal. Administration was seen as more burdensome than that in national programmes, and – except where networking and scale are important – participants prefer to use national programmes. The FPs therefore focus in areas where they generate 'European Added Value'. Projects are mostly 'additional' in the sense that they would not have been conducted without European funding. If attracting SME participation continues to be a policy objective, greater parts of the Framework should explicitly address their needs, which currently appear poorly handled.

Framework projects primarily produce knowledge and networks, strengthening European-level human capital and RTD capabilities across borders. Their role is therefore quite distinct from nationally funded projects. There is evidence that the Framework Programmes have positive effects on the behaviour of the research community, competitivity, jobs, regulation and the environment.

Member states need to improve national RTD strategies by taking better account of European-level programmes and policies.

The Five Year Assessment in 2000 concluded that the FP was not on its own sufficient to reach the Lisbon objectives. There was a need for a 'real European RTD policy', involving a more flexible FP and the incorporation of

other instruments and actions. Since then, integration between FP design and wider EU policy for innovation and R&D appears to have improved, though there is further to go in this direction.

The intervention logic that connects the high-level and operational goals of the FP is poorly articulated, making an overall evaluation of the FP difficult. The Framework needs more systematic planning, clearer objectives and a stronger link to an evidence base. This would ease evaluation and, arguably, improve FP performance. Framework evaluation needs to be more clearly linked to understanding, diagnosis and wider policy for the European Research and Innovation System. This implies a more systemic approach to evaluation and to related studies needed to support policy development and planning, new and additional evaluation methods and further improvements in the organisation of evaluation to link it better to an improved planning and policy framework for EU R&D.

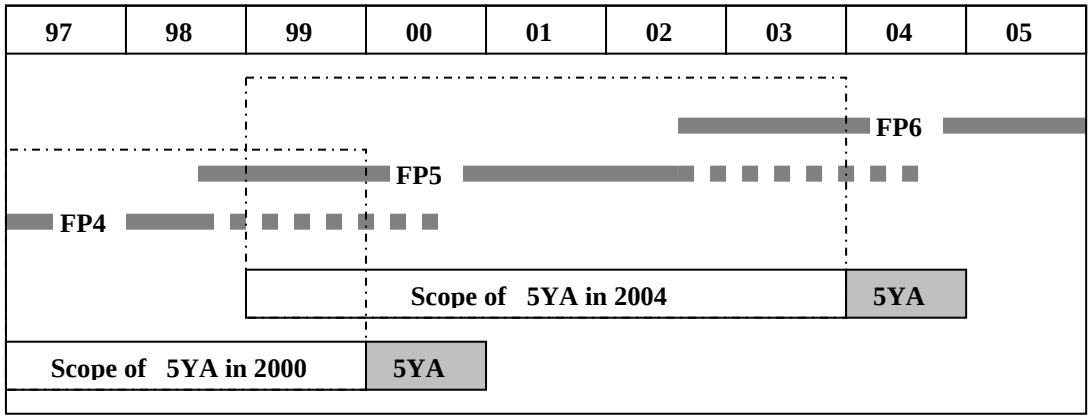
What the Evaluation Record tells us about Framework Programme Performance[1]

1 Introduction

This paper reviews the record of programme evaluation of the Framework Programmes in the period 1999 to mid-2004 and distils key messages from it. The analysis served as an input to the Five-Year Assessment of the EU RTD Framework Programme (FP) 1999-2003.

The Five-Year Assessment is a mandatory exercise foreseen in the Decisions of the Council of Ministers concerning the Framework Programmes. Given the requirement that such an exercise take place before the Commission submits its proposals for a new Framework Programme, it can be seen as part of the programming cycle for EU R&D programmes. **Exhibit 1** shows that the Framework Programmes are in principle four years long. However, because the programmes launch projects that can last several years, there is always an overhang of projects at the end of each programme (illustrated by the dotted lines). The Five-Year Assessments (5YA) are made every fourth year, but have a **scope** that covers the previous five, as the dashed rectangles in the **Exhibit** illustrate. The timing of the Five Year Assessments (5YAs) therefore means that they combine an ex post evaluation of the previous programme and a mid-term appraisal of the current one in order to provide recommendations for the coming FP.

Exhibit 1 Timing of Framework Programmes and Five-Year Assessments



Five Year Assessment panels evaluated the FPs in 2000 and 2004. However, the detail of the way the Framework Programmes are evaluated has changed during the period considered. In 2000, the 5YA panel was the apex of an evaluative pyramid. Other panels evaluated the individual Specific Programmes (SPs – the major blocks into which the FPs are divided). These

[1] This paper has benefited from inputs and comments from John Clark, Alessandro Muscio and James Stroyan, which I gratefully acknowledge. The normal disclaimer, of course, applies. This document should not be read as necessarily reflecting the position of the European Commission or its staff.

SP evaluations were the main inputs to the Five Year Assessment panel's work. The 2004 5YA dispensed with SP-level panels (though there was a separate panel evaluation of the IST Programme). Four studies[2] and four analyses by independent experts were commissioned to provide additional inputs to the 5YA panel, of which this is one, but the 2004 panel essentially had to tackle all the evidence itself, without the support of contributory panels.

Another important part of the FP level evaluation set up is provided by the Annual Monitoring exercise. Up to 2003, in addition this consisted of SP-level panels plus panels to cover the FP (and more recently the European Research Area – ERA) . From 2004, these were replaced by a single panel, providing annual monitoring of the whole of the Framework Programme. The monitoring and evaluation structures have therefore been simplified, with the overall panels now having to assemble evidence about the entire FP, rather than being able to combine results from lower level panels.

Framework Programme level evaluation is the responsibility of DG Research. Following a re-organisation at the start of 2004, DG Research now has two separate evaluation units at the strategic level, both located within the Directorate for Co-ordination of Community Activities. The Unit for *'Planning, Programming and Evaluation'* handles overall coordination of evaluation, planning and system design, as well as the execution of horizontal and strategic level ex post and intermediate evaluations. This joint responsibility for both evaluation and planning and programming activities reflects the linkage foreseen in recent Commission reforms. However, the coordination of ex ante evaluation, specifically the preparation of an ex ante Impact Assessment of proposals for the next Framework Programme, is carried out separately by the Unit *'Analysis of the impact of Community Actions'*. Evaluation units or functions are also located in other DGs with responsibility for managing parts of the Framework Programme, notably DG INFSO which has a substantial programme of evaluation work and during 2004 carried out its own Five Year Assessment of the IST area.

Evaluation is also carried out at the level of the SPs and individual research activities. The approach differs between DGs. In DG Research such evaluation work is the responsibility of the operational units concerned, with

[2] The specific studies that were implemented to provide input for the Five-Year Assessment (1999-2003) were as follows: an impact evaluation of Framework Programmes 3 and 4; an impact evaluation of Framework Programme 5; a bibliometrics study of scientific publishing resulting from Framework Programmes 4 and 5; a study of High Impact Research Activities under Framework Programmes 4 and 5. Additional work of an evaluative nature was carried out through four analyses by independent experts, the subjects being as follows: an analysis of Europe's changing economic landscape; an analysis of Europe's changing research policy landscape; an analysis of the implementation of Community research; an analysis to synthesise the record of evaluation reports and studies on Community research at EC and Member States level. The relevant reports from this work are available at:
http://forum.europa.eu.int/Public/irc/rtd/fiveyearasskb/library?l=/&vm=detailed&sb= Title.

coordination[3] provided at the DG level through the central evaluation units although in other DGs, such as DG INFSO, the operational level evaluation is carried by the central evaluation functions.

Finally, evaluations of FP research impact and especially the participation in FP research are carried out at the Member State level.

Over recent years there have been significant moves to improve coordination and reporting on evaluation work. First, the introduction of the Multiannual Evaluation Plan as a part of the DGs Annual Management Plan has been an important step towards greater clarity and a more strategic approach. Second, the Annual Evaluation Review that is produced by DG Budget and covers the whole of the Commission, has been developed to allow for detailed reporting on completed and ongoing evaluations. For the 2004 Five Year Assessment a database was prepared covering past evaluations of FP research. Although some research areas have not been subject to an individual evaluation during the period covered it nonetheless provides an important resource and was the basis for this paper.

The available evaluations are of four types

- The panel reports from the last Five-Year Assessment, in 2000
- FP-wide impact studies
- Evaluations and impact studies at Specific Programme (SPs) and research activity level within the Framework Programme or on thematic issues, such as socio-economic aspects
- Studies, which evaluate impacts of the FP at national level

In what follows, we begin by discussing goals and planning in the FP. We briefly review the reports of the last Five-Year Assessment then go on to discuss the methods used in the main body of externally performed evaluation reports. We next summarise and generalise the findings of those evaluations. In the following section, we raise issues about the practice of evaluation of the FPs. Our paper ends with some suggestions for improving the evidence base, the evaluations and the use of evaluation in the broader process of policy analysis and development.

This review of evaluations has a number of limitations. As already noted, the evaluations we have obtained and read do not provide complete coverage of the FPs and those that are available were not designed to provide a basis for this type of meta-evaluation. In some areas we have had to extrapolate from a limited amount of evaluative material – using our knowledge of R&D programmes and their evaluations in many countries – in order to reach conclusions. Much of what we have to say is therefore suggestive rather than conclusive. However, we have tried to be clear about

[3] An important step has been the introduction of the Annual Evaluation Plan that is part of the Annual Management Plan for the DG.

sources throughout, so that the reader can herself assess the degree of support there is for our interpretation in the underlying body of evaluation material.

2 Goals and Planning in the Framework Programmes

Good planning and good evaluation go hand in hand. The clearer you are about exactly what you are trying to achieve and how you are going to achieve it, the more likely you are to reach your goals and the easier it is to find out whether and when you have done so.

The overall objective of the FPs laid down in the Treaty establishing the EU – namely, strengthening the scientific and technological bases of European industry and contributing to the quality of life of its citizens – is extremely permissive. It creates a legal basis for the EU and its executive to operate a wide range of RTD funding activities. This is a helpful freedom that needs to be complemented by specific analysis and planning in order to achieve specific ends.

Internationally, there are no absolute standards for planning major R&D programmes. Practices are everywhere rather mixed and rigour seems increasingly difficult to achieve the higher one ascends a pyramid of interrelated programmes. However, generally accepted principles include:

- The need to link programme planning to an analysis of needs and problems
- Limiting the scope of interventions to those where the state has a legitimate role (for example, because there is a market failure or a systems failure preventing the solution of problems) and a legal basis for action
- Identifying the mechanisms through which it is reasonable to expect the planned programme to lead to the intended goals (In the Commission, this tends to be called an 'intervention logic'. Elsewhere, it is sometimes also called a 'programme theory')
- Making clear statements of goals, which need to be non-trivial and testable[4]
- Ensuring that the intended programme or actions are sufficient to reach the goals
- Monitoring progress and evaluating results, in order to create a feedback loop to policy making and future planning, in addition to providing accountability to the taxpayer

Techniques exist for programme planning. The Commission promotes the use of Logical Framework Analysis in this kind of planning context. Logical frameworks are not the only useful planning tools. For example, logic

[4] In fact, there is fairly general agreement that goals should be SMART: Specific, Measurable, Attainable, Realistic and Timely

diagrams and modelling can also be used[5]. Programmes can contain sub-programmes, which contain projects, which contain sub-projects, and so on. However, a programme is not simply a higher-level project. Rather, it is the 'hinge' or transformation mechanism that brings together technical activities (projects) in order to achieve socio-economic goals. This means that programme design and management is more than project management: it involves wider skills and interrelationships. Work is in progress[6] to use developments of logic modelling in the monitoring and evaluation of the IST programme, but there is no evidence in the existing evaluation record of similar techniques being used at the planning stage of past FPs.

The FPs have been in almost constant change since they were first conceived, which complicates evaluation. FP5 (1998 – 2002) represented a new direction compared with the earlier FPs, shifting significantly towards more specific applications-relevant areas of science and technology. The size and allocation of the budget among different subjects changed following extensive consultation within and outside the Commission, but no rationale for this is discussed in the legislation or the lower-level documents that together make up the nearest thing the FPs have to programming documents.

Analysis of FP5 suggests that the activities and goals defined in research terms are broadly consistent with higher-level socioeconomic goals. The goal hierarchy is generally specified in terms of research and knowledge. However - at least in terms of the available documentation there is no explicit link to the amount of resources needed or the actors who should be involved, or their capabilities. It is not clear what assumptions are being made about factors that may be vital to making good social use of the research to be funded, such as framework conditions. At all levels, goals are specified in an undemanding way – mostly to 'do activities' within some specified area, rather than to produce specified outcomes.

Technical goals at the bottom of the goals hierarchy are more clearly defined than socioeconomic goals at higher levels but the 'intervention logic' that should connect the two is not explicit. (This problem of a 'missing middle' between low-level technical activities and high-level socioeconomic goals is not unique to the FPs: it is very common in R&D programming.) In practice, the legislation and the implementation of the FPs rely heavily on the use of selection criteria to achieve consistency between low-level activities and high-level policy goals. But putting in place processes that make projects and programmes **consistent** with policy and ensuring that they are

[5] Elsewhere, in work for the Commission, we have shown how to connect this style of planning into RTD programme design and management more generally. Erik Arnold, Patries Boekholt and Patrick Keene, *Some Good Ideas in RTD Programme Management*, report to the VALUE programme, Brighton: Technopolis, 1996

[6] Technopolis, *Developing a methodological framework for high-quality assessment of the IST-RTD effects (results and impacts) at the 'Strategic Objective' level*, Commission contract No 29000, Revised inception report and detailed methodology, Brighton: Technopolis, December 2004

sufficient to achieve the policy aims are by no means the same thing. The Framework Programme **is** a framework: it is not a plan. While history teaches that there are limits to planning and that all good plans are flexible, it is reasonable to argue that achieving specific goals requires rather specific plans.

FP5 appeared rather isolated from other aspects of EU policy. However, the integration between research and non-research policies has increased during its lifetime and at the start of FP6 by a series of policy initiatives: the European Research Area (ERA) project; the consequences of the Lisbon declaration; and the definition of the Barcelona goal. These increasingly require consistency and connection between R&D and other policy responsibilities of the Commission, intervening at multiple points in the European Research and Innovation System (ERIS). In the past 4 years, the Commission has much more evidently been using **strategic intelligence** in the form of studies, panels, benchmarking and so on as an 'objective' complement to the more traditional process of stakeholder consultation about research. This clearer connection between policy, analysis and evidence has been supported by the decision to adopt an 'open method of co-ordination' of research and innovation policy as a contribution towards creating the ERA. Policy is beginning to tackle higher-level questions about the ERIS as a whole, especially those of critical mass, and increasingly it is becoming possible to say that the Union has the beginnings of a research and innovation policy.

While the formulation of goals and the intervention logic in FP6 remain (from a 'pure' planning perspective) problematic, the Framework Programme is now being brought into closer alignment with other policies. New instruments were introduced in FP6, aiming to build European-level critical mass within the research-performing community. Seven thematic areas were defined, partly with the idea that focus would help build this critical mass, though there is no explicit rationale given for the choice of these, rather than other, areas or an analysis that links the achievement of sub-goals to overall goals. The FP6 legislation appears as a still frame from the moving picture of policy development. It seems that a better integration of RTD policy and FP evaluation will be needed.

3 What emerged from the 2000 Five-Year Assessment?

The main conclusion of the 2000 Five-Year Assessment (5YA) was that the FP was insufficient as the means to reach the Lisbon objectives. There was a need for a 'real European RTD policy', involving a more flexible FP and the incorporation of other instruments and actions.

I reviewed the sub-panel reports for the 2000 5YA exercise. The work of the sub-panels was informed by a Commission survey carried out in late 1999 concerning finished projects in FP3 and FP4. Individual sub-panels

supplemented this information, using their own data-gathering procedures and earlier reports.

The discussion in the reports is understandably more sector-specific than that in overall FP or national evaluations, with reference to low-level programme objectives and attributes, but the broad conclusions tend not to be very different. Conclusions on broad objectives such as[7] "It is clear that FP3/4 has definitely contributed to European cohesion and competitiveness" tend to be based on beneficiaries' and officials' general satisfaction with programmes rather than hard data or systematic analysis of impacts. More interesting, perhaps, is the apparently widespread concern with the objectives themselves – that they are too broad, poorly specified, and there are too many of them. In several cases, the panels criticised inadequate strategic thinking in programme design and inadequate synergies with national and other EU programmes. The following (from the 'Competitive and Sustainable Growth') report is fairly typical:

> … a more systematic formulation of objectives is recommended…there might be… an increasingly difficult trade-off between the different objectives (science and technology excellence, industrial competitiveness, broader socio-economic goals, contributions to European cohesion, SME participation). The concepts of 'European Dimension' and 'European Added Value' have to be further clarified and translated into programme objectives and corresponding programme selection criteria….

On ERA, there is some optimism that progress has been made, while the position of SMEs is almost universally seen as problematic. Panels are generally positive and endorse the Specific Programmes they evaluate. They regard scientific quality as high and are positive about 'intermediate' or first-order outputs: knowledge/skills acquisition and transfer; networking; and researcher mobility. They point to problems relating to the complexity of application forms and procedures and express disappointment about the level of SME involvement. They tend to argue that the moves towards more market-oriented, less risky projects risk neglecting longer-term, basic research are problematic. Dissemination and exploitation of results could be improved in a number of cases.

There was overwhelming support in the 2000 5YA for the continuation of Community research activities and for a very substantial increase in funding. The Commission's formal responses to this exercise overwhelmingly endorse the findings of the Assessment panels, with no significant disagreements. Very broadly, the findings are 'noted with interest' or 'warmly welcomed', the

[7] Alain Pompidou et al, *Five year assessment report related to the specific programme: User-Friendly Information Society, covering the period 1995-1999*, Brussels, European Commission, 2000

latter often implying that a proposal had, already, or was being, implemented.

4 The Evaluations we Studied and the Methods they Used

The evaluations reviewed for this paper, are listed in **Exhibit 2**. (The **Exhibit** also provides a systematic shorthand for referring to the studies, which are used in the analysis of what the evaluations say, below.) Some 'evaluations' in the DG Research database are simply national analyses of participation data and are descriptive rather than evaluative, so they have not been included in the survey.

The methods used in the evaluations are summarised in **Exhibit 3**. Individual techniques have limitations, so it is important for evaluations to use multiple techniques in parallel, in order to increase their reliability.

A key technique used in the evaluations studied is sending questionnaires to project participants, normally using a variant of the Alvey questionnaire[8], and this is generally complemented by the use of other methods in order to permit triangulation. (The EU level impact studies are intended to support broader 5YA inquiries, so it would be wrong to criticise them for involving only one method.) Control groups are rarely used (Austria and Denmark are exceptions) – often because of budget pressure on the evaluations. The Swedish study that intended to discover whether FP research had lower quality than nationally funded research is an outlier, in that it uses participation and other data from the national research council to benchmark the standing of Swedish researchers. It is noteworthy that the studies do not make use of baselines – indeed baselines are not drawn up, though this is admittedly difficult to do in the context of R&D programmes. This means there is no 'before' picture to compare with the situation 'after' the intervention.

Overall, the picture is of rather homogenous methods – a homogeneity that probably ought to decline as evaluations tackle higher-level, more systemic and more policy-oriented questions in future.

[8] Developed by SPRU and PREST during the evaluation of the Alvey Programme for advanced industrial IT research in the UK during the early 1980s, this questionnaire has become pervasive in European R&D programme evaluations at national and international levels. See Ken Guy, Luke Georghiou, Paul Quintas, Hugh Cameron, Michael Hobday and Tim Ray, *Evaluation of the Alvey Programme for Advanced Information Technology*, London: HMSO, 1991

Exhibit 2 Evaluations Surveyed in this Section of the Report

	FP4	FP5	FP6
EU / Framework	FP3-4 Impact[9]	FP5 Impact[10]	
SP / Action	FP4 Impact BioMed2[11] FP4 Impact BRITE-EURAM[12] FP4 Impact Growth[13] FP4 Impact International RTD Co-operation[14] FP4 Impact studies of Joule, Thermie and NNE[15] FP4 Impact Telematics[16]	FP5 Mid-term Quality of Life Action 6 (Ageing)[17] FP5 Impact of dissemination in Environment[18] FP5 Gender Impact IST[19] FP5 Impact Genomics[20]	
Thematic / Horizontal		FP5 Socio-economic dimension[21]	Marimon Report[22]
National	FP4 Impact Austria[23] FP4 Impact Denmark[24] FP4 Impact Germany[25] FP4 Impact Ireland[26] FP4 Participation Finland[27] FP4 scientific quality of Swedish participation[28]	FP5 Impact Norway[29] FP5 Impact Finnish Universities[30] FP5 Participation Finland[31]	FP4-6 Impact UK[32]

[9] Decisia, HLP Developpement and Euroquality, *Assessment of the Impact of the Actions completed under the 3rd and 4th Community Framework Programmes for Research*, Levallois-Perret: Decisia 2004

[10] Atlantis, Wise Guys and Joanneum Research, *Assessment of the Impact of the Actions completed under the 5th Community Research Framework Programme (1999-2003)*, Work in Progress Report (unpublished), 2004

[11] EC, *Impact Assessment of the Biomedical and Health Research Programme BIOMED2 (1994-98)*, Brussels: EC, 2002

[12] EC, *BRITE-EURAM, Making a Lasting Impression on Europe*, Brussels: EC, 2002

[13] GOPA Consortium, *Impact Assessment of Finished Projects of the EC Research Programmes in the Fields Covered by the Present Growth Programme*, Bad Homburg: GOPA, 2003; GOPA Consortium, Evaluation of Finished Projects in the Fields Covered by the Pesent Growth Programme, Bad Homburg: GOPA, 2003

[14] EC, *Impact Assessment Report on the Specific Programme: International RTD Co-operation, 4the Framework Programme (1994-1998)*, Brussels: 2003

[15] EC, Clean, *Safe and Efficient Energy for Europe, Impact Assessment of Non-Nuclear Energy Projects Implemented under the Fourth framework Programme, Synthesis Report*, EUR 20876/2, Brussels: EC, 2003. This summarised impact studies in the area, whose contents we also include under this reference

[16] Gabriella Cattaneo TAP-ASSESS Consortium, *Telematics Projects Socio-Economic Impact Assessment: Conclusions and Recommendations*, presentation to the EC, Brussels, 2000

[17] EC, Mid-Term *Assessment Report, Key Action 6, The Ageing Population and Disabilities, 1999-2002*, Brussels: EC, 2003

[18] EC, *Impact Study of Result Dissemination in the Field of Environment and Sustainable Development, Brussels*: EC D-G Research, 2003

[19] European Commission, *Gender Impact Assessment of the Specific Programmes of Framework Programme 5: User-Friendly Information Society (IST)*, Brussels: European Commission, 2001

[20] EC, *Analysis of genomic research supported under FP5, Quality of life and management of living resources (1998-2002)*, EUR 20811, Brussels: DG-Research, 2004

[21] EC, *The Overall Socio-Economic Dimension of Community Research in the Fifth European Framework Programme*, Brussels: EC, 2003

[22] *Evaluation of the Effectiveness of the New Instruments of Framework VI*, Report of a High-Level Expert Panel chaired by Ramon Marimon, Brussels: EC, 2004

[23] Andreas Schibany, Leonhard Joerg, Helmut Gassler, Katharina Warta, Dorothea Sturn, Wolfgang Polt, Gerhard Streicher, Terttu Luukkonen and Erik Arnold, *Evaluation of Austrian Participation in the 4th EU Framework Programme*, Technopolis Ltd, Joanneum Research and VTT, Vienna: BMVIT, 2001

[24] Ebbe K Graversen and Karen Siune, *Danish Research Co-operation in EU: Extent, Return and Participation, An analysis of co-operation in the 4th EU Framework Programme*, Report 2000/7, Århus: Danish Institute for Studies in Research and Research Policy, 2000

[25] ISG, *Europäische Forschungsrahmenprogramme in Deutschland*, Köln: 2001

[26] Ken Guy, Jane Tebbutt and James Stroyan, The Fourth Framework Programme in Ireland: An Evaluation of the Operation and Impacts in Ireland of the EU's Fourth Framework Programme for Research and Development, report to Forfás by Technopolis, Dublin: Forfás, 2000

[27] Terttu Luukkonen, Sasu Kalikka, Pirjo Niskanen and Riika Eela, *Finnish Participation in the Fourth Framework Programme*, VTT Technology Studies report 1999/4, Helsinki: TEKES, 1999

[28] Royal Swedish Academies of Sciences and Engineering, *Qualitative Aspects of Swedish Participation in EU Research Programmes*, Stockholm: Dokumenta report No 66, 1999

[29] NIFU, STEP and Technopolis, *Evaluation of Norway's Participation in the EU's 5th Framework Programme*, Oslo: STEP, 2004

[30] Pirjo Niskanen, *Finnish Universities and the EU Framework Programme – Towards a New Phase*, VTT Technology Studies, Helsinki: VTT, 2001

[31] Marjo Uotila, Pirjo Kutinlahti, Soile Kuitinen and Torsti Loikkanen, *Finnish Participation in the EU Fifth Framework Programme and Beyond*, VTT Technology Studies, Helsinki: Finnish Secretariat for EU R&D, 2004

[32] Paul Simmonds, James Stroyan, John Clark and Ben Thuriaux, *The Impact of the Framework Programmes in the UK*, London: Office of Science and Technology, 2004

Exhibit 3 Methods Used in Evaluations Summarised

Study	Scientific Peer review	Expert / consult-ant review	Particip-ant Q'aire	Inter-views	EC data / docu-ment review	Control group	National research council data
EU level							
FP3/4 Impact			X				
FP5 Impact			X				
SP / Action Level							
FP4 Impact BioMed2	X		X		X		
FP4 Impact BRITE-EURAM		X	X		X		
FP4 Impact Growth		X	X	X	X		
FP4 Impact International RTD Co-operation		X	X	X	X		
FP4 Impact Joule/Thermie/NNE		X	X	X	X		
FP4 Impact NNE		X	X	X	X		
FP4 Impact Telematics (TAP-ASSESS)		X	X		X		
FP5 Mid-term QoLife Action 6 (Ageing)	X			X	X		
FP4 Impact Growth		X	X	X	X		
FP5 Impact of dissemination in Environment	X				X		
FP5 Impact Gender IST				X	X		
FP5 Impact Genomics	X		X	X	X		
Thematic							
FP5 Socio-economic dimension				X	X		
Marimon Report			X	X	X		
National							
FP4 Impact Austria			X	X	X	X	
FP4 Impact Denmark			X	X	X	X	
FP4 Impact Germany			X	X	X		
FP4 Impact Ireland			X	X	X		
FP4 Participation Finland					X		
FP4 Quality of Swedish participation			X	X			X
FP5 Impact Norway			X	X	X		
FP5 Impact Finland Universities			X	X			
FP4-6 Impact UK			X	X	X		

5 What the evaluations say

Generically, evaluations of policy need to answer four questions in order to provide accountability and feedback that can be used in learning

- Are we doing the right thing? (Appropriateness)
- Are we doing it well? (Quality and efficiency of implementation)
- What happens as a result? (Effects, impacts)
- What should we do next? (Feedback to improve appropriateness and implementation)

For the period in question and as noted above, the evaluation studies being considered do not provide a complete coverage of the FP. Its not therefore possible to use an arithmetically based method to derive findings at the level of the FP, using arguments of the form '15 evaluations say A and 5 say B, therefore A'. Rather, an attempt has been made to interpret and generalise the results of individual evaluations where this appears to be reasonable.

5.1 Appropriateness

The Five Year Assessment process tackles the question of appropriateness of the FP as a whole. Very little more can be said about the current or continuing appropriateness of the FPs on the basis of other evaluations, which tend not to state, discuss or analyse the higher-level goals of the FP or of individual SPs. Goals are more usually discussed at the level of action lines or clusters of projects, where they are generally stated in terms of performing a particular category of research. Evaluators tend to have to 'go fishing', casting their nets wide and looking for a range of possible effects rather than being able to test focused hypotheses about outcomes and impacts. In view of this, we were surprised how rarely goals were criticised. An example of an exception was the Non-Nuclear Energy programme (NNE), which was said[33] to lack identity and direction. It needed more strategic intelligence in order to identify the specific goals needed to enable the programme to make a difference.

The FP4 Impact Austria study was methodologically more ambitious than most, and attempted an evaluation of the 'European paradox' justification for FP4. Unsurprisingly, in a study that was primarily resourced to understand the costs and benefits of Austrian participation in the FP, this was somewhat inconclusive, beyond echoing the 5YA finding that the FP is far too small an instrument to correct that paradox. On the other hand, if the FP was essentially about creating common European strengths through networking, it was an adequate instrument.

More technically-oriented evaluations focus on the technical goals of the work programme. The typical goal-setting mechanism of "extensive consultation with sector actors"[34] does not explicitly connect technical goals to the goals of the FP as a whole. Technically focused evaluations, however, do discuss the need to update technical goals. Since these exercises tend to be done by members of the beneficiary communities, they play not only an evaluative role but also act as an extension to the wider process of stakeholder consultation. For example, FP4 Impact Biomed2 pointed out the lack of 'translational research' funded in the programme and recommended that this should be included. Similarly, the FP5 Impact Genomics objected that bioinformatics was not adequately handled in the Quality of Life programme of FP5 and that it should be included in future calls.

[33] FP4 Impact NNE
[34] FP5 Mid-term Quality of Life Action 6 Ageing

National impact studies tend to criticise the national ability to develop a coherent R&D strategy and to connect that with the formulation of the FPs. Mutual influence of FP and national R&D strategies is seen as very limited.

5.2 Implementation

The evaluation record implies that the scientific quality of FP work is broadly good. The FP has strong proposal assessment and project monitoring procedures that help ensure that the quality of the projects funded is high, in scientific and technological terms. However, few evaluations try to measure quality directly.

The FP4 Impact Biomed2 assessment panel criticised scientific quality in the programme. In some sub-areas, only 12-15% of projects published in the top 10% of journals. Some 25-50% of projects were high quality, and the panel drew the conclusion that the proposal assessment process was inadequate. The NNE programme was criticised[35] for having too high a share of failing projects, which tended to duplicate other work, were too far from the technological leading edge or lacked the needed human or financial resources to reach their goals. But these are outliers: the other seven evaluations reviewed that tackle quality issues found little to criticise.

Members of the 'basic' research community in at least some countries have argued that the scientific quality of FP work is low. Rather than being based on evidence from projects, however, this tends to amount to a view that much of the FP research is rather applied, combined with an un-stated assumption that only 'basic' or 'researcher-initiated' research free of social and industrial influence can be can be good research.

After some years of claims at a national level that FP projects had low scientific quality, the Swedish academies of science and engineering ran an inquiry[36] in 1999, which found that the EC has a very high-quality proposal assessment process, at least as good as Swedish national practice. 40% of respondents said the FPs had no effect on scientific quality, while 42% believed it raised quality. The study seems largely to have silenced criticism in Sweden about FPs funding poor quality work. In contrast to the normal evaluation finding concerning FP projects that knowledge production is a key output, the FP5 Impact Finnish Universities study[37] nonetheless found that, in the eyes of the Finnish university community, "new **scientific** knowledge is still quite rarely an important benefit of the projects." Less than half the respondents considered EU-funded research projects as internationally high standard, whereas 17% thought they are internationally low standard. Compared with their own research, Finnish university respondents saw FP research as less innovative, more short term, more application-oriented and technically less complex. Of those who claimed to be doing basic research in

[35] FP4 Impact NNE
[36] FP4 Scientific Quality of Swedish Participation
[37] FP5 Impact Finnish Universities

an FP project, 73% saw the quality of their own project as of an internationally high standard. Only 36% of those doing applied work in the FP saw it as of a high standard. The majority wanted to see more basic research funded by the FP. More conceptual work on scientific quality and perceptions of scientific quality would be useful, to help clarify the debate about quality in programmes like the FPs, which have a largely socio-economic purpose.

FP funding, like R&D funding in many OECD countries, has tended to shift in recent time slightly towards more 'basic' or researcher-initiated research, through the Networks of Excellence of FP6 and its New and Emerging Science and Technology (NEST) activity. The rationale for this shift is not clear from available evaluations, but appears to reflect a wider shift in the influence of the basic research community. Some evidence also suggests that the balance of participation between those in the knowledge infrastructure (universities, colleges and research institutes) and those in industry has shifted in favour of the former in the transition to FP6, partly because the new Networks of Excellence instrument do not involve many companies and partly because work programmes have shifted away from the strong applications focus of FP5[38]. This balance between participation by the knowledge infrastructure and industry has always varied by country, however, presumably as a function of national structures and specialisation. **Exhibit 4** shows the proportions of project participations by different categories of actor in a number of countries, according to national impact studies.

Exhibit 4 Project Participations in FPs

	HEIs	RI	SMEs	Large firms	Other
FP5 Impact	28%	25%	32%	8%	7%
FP4 Impact Austria	31%	15%	29%	10%	15%
FP5 Participation Finland	33%	31%	16%	11%	9%
FP5 Impact Norway	24%	37%	32%		7%
FP4 Impact Ireland	45%	12%	32%		11%
FP5 Impact UK	46%	17%*	27%		10%

(*) Note: figures also include private commercial research centres. However the total proportion of such institutions amount to a mere 2-3%.

It appears that there is still a wider industrial constituency, which the FPs could reach. During FP4, only some 40% of research-active German companies and 50% of the biggest 500 firms in Germany that also received subsidy from the BMBF participated in the FP[39]. In the UK there were 417 fewer UK organisations involved in FP5 than in FP4 - a 14% drop overall, though the proportion of company participations remained stable. Many of the UK's major companies said they were not participating because they felt the benefits merely balanced, or were outweighed by, the costs. Some large

[38] Marimon Report
[39] FP4 Impact Germany

Danish firms stayed out of the FP for strategic reasons, not wanting to share knowledge with key competitors[40].

Many organisations' participation in FPs is short lived, but there is a core of frequent participants, who sit at the heart of evolving networks. Evaluation and study[41] evidence supports the idea that there are virtuous circles that lead a limited sub-set of participants to become major actors within the FPs. In the 2004 FP5 impact survey[42] 55% of the FP5 participants also participated in FP6 while more than 70% applied. Around 60% of the research institutes, universities and large firms that participated in FP5 continued their participation in FP6, but this percentage dropped to around 35% for SMEs.

FP4-6 Impact UK investigated the numbers of UK participations in FP4 and FP5. The analysis of the results of the survey confirms that the bulk of the UK's participation in Framework is accounted for by a relatively small number of organisations that (a) get involved in successive Framework Programmes and (b) tend to undertake a large number of projects each. In contrast, the vast majority of UK participants take part in just one programme and in an average of about one and a third projects.

Exhibit 5 Continuity of involvement from FP4 to FP5[43]

	Took part in FP4 only	Took part in FP5 only	Took part in both FP4 & FP5	All combined
Number of organisations	2,275	1,859	774	**4,908**
Number of participations	3,121	2,471	16,556	**22,148**
Proportion of all UK FP4/5 participations	14%	11%	75%	**100%**
Average number of participations per org.	1.37	1.33	21.39	**4.51**

Source: EC data – Technopolis analysis

Other evidence supports the idea that there are mechanisms that lead to a limited sub-set of participants becoming major actors within the programmes. "Earlier experience of international research collaboration is one of the most important factors facilitating the activity … in EU research programmes. Expectedly, the EU research programmes attracted especially university researchers who had earlier collaborated with firms." 80% of participants said they would apply again to the programme, while only 50% of non-participants intended to do so[44]. In BRITE-EURAM, half the companies were in their first such collaboration, while two thirds of the RTOs

[40] FP4 Impact Denmark
[41] Sefano Breschi and Lucia Cuismano, 'Unveiling the texture of a European Research Area: Emergence of oligarchic networks under EU Framework Programmes,' paper presented at the conference *Evaluation of Government Funded R&D Activities*, Vienna, May 15-16, 2003
[42] FP5 Impact
[43] FP4-6 Impact UK
[44] FP5 Impact Finnish Universities

had experience of multiple collaborations[45]. Bringing new partners to various sources of funding appears to be core business for RTOs, something they do on a national level as well.

A factor promoting stability among a core of frequent participators is the fact that (like other network R&D programmes) the FP does not generate wholly new R&D networks, but causes **network extension**. Evaluations of network R&D tend to find that R&D networks evolve over time, rather than being newly constructed for each funding opportunity[46]. The FP4 Impact Denmark study found that networks were based on existing networks but that 82% of companies and 90% of universities also established new international research partnerships through the projects. The Finnish university impact study[47] found that EU collaboration did not crowd out academics' other international networks. Rather, it led to an increase in participants' non-FP international networking. So a positive aspect of having a stable core of participants is that its composition evolves, and that – through established networks – new participants can sample participation, even if the majority then decides not to repeat the experience.

For the knowledge infrastructure, the FP is but one among a number of sources of routine project funding income. It seems that *pecunia non olet* and the Commission's money is as gratefully accepted as any other coin. Given the chance to get money on less demanding conditions, of course, that chance will be taken – a finding as clear in FP evaluations as in evaluations of national programmes. Other things being equal, participants prefer national programmes to the FP, but they recognise the need to go to the FP (or an equivalent) where international networking, socio-political objectives and exploitation are important to them[48]. Austrian university FP4 participants obtained funding from multiple sources – the FP was simply one part of a bigger funding portfolio – while 65% of Danish company participants were involved in other international R&D co-operations[49]. "Especially interviewed heads of units perceived EU collaboration as an important channel through which to obtain research funding for their units and to gain prestige. The participation of universities depends on the availability of other research funding from national or other international sources. A decline in the research funding is likely to increase researchers' interest in seeking EU funding, whereas the availability of other funding may decrease its relative attractiveness."[50]

[45] FP4 Impact Growth

[46] See for example Sven Faugert, Erik Arnold, Alasdair Reid, Annelie Eriksson, Tommy Jansson and Rapela Zaman, Evaluation of the Öresund Contracts for Cross-Border R&D Co-operation between Denmark and Sweden, Stockolm: VINNOVA, 2004; Sven Faugert, Erik Arnold, Ben Thuriaux and Bo Sandberg, *NUTEKs program VAMP: en utvärdering av programstrategi, genomförandet och resultat*, SIPU Utvärdering och Technopolis, Stockholm: NUTEK, 2000

[47] FP5 Impact Finnish Universities

[48] FP4 Impact Germany

[49] FP4 Impact Austria; FP4 Impact Denmark

[50] FP5 Impact Finnish Universities

This reminds us that not only do research performers compete with each other to obtain funding, but that providers are also in competition among themselves. Participants often have their own 'real projects', which they sub-divide and package in various ways to make them attractive to funders. Funders easily fall victim to the 'project fallacy'[51]: the idea that what they define in **administrative** terms as a project also has a wider meaning for the research performers. Correspondingly, aside from the frequent need to build communities and networks around R&D programmes, there is little reason from the programme manager's perspective to capture participants' 'hearts and minds', provided they produce outputs and outcomes that promote programme goals. Whether these results also support participants' 'real projects' need not be a matter of concern.

Knowledge Infrastructure participants attach much higher importance to FP participation than do industrial participants. For them, the FP is a source of operating revenue[52]. For companies, participation is a means to other ends. Unlike members of the Knowledge Infrastructure, companies tended to regard the FP as a more marginal source of funding[53].

Overall, the main motives behind the involvement of research institutions in FP projects appear to be the opportunity to establish new networks, the possibility to access research funding and the opportunity to achieve knowledge-oriented objectives. As FP5 Impact Finnish Universities suggests, this is not necessarily the same as **scientific** knowledge.) Industry participants seem to be involved in FP projects primarily in order to achieve knowledge-oriented goals and access financial resources. While participants in most SPs are involved in order to reap intermediate or knowledge outputs, in certain more commercial output oriented programmes such as BRITE-EURAM and NNE, participants' goals are much more short term and commercial. Thus 70% of companies in the NNE programme developed a new products or processes[54]. The FP5 Impact Growth study argues that homogenous instruments can conceal very different types of activity that – by implication – might better be addressed using different, dedicated instruments. Thus, the study found that there were two groups of successful projects. One was moderately ambitious, close to market, risky and product-oriented. The other was technically complex, high risk, long term and involved a high proportion of large companies.

Administration is a sore point. It seems to be axiomatic among academic research performers that administration is unnecessary and that funders should simply give them money to do their research. The evaluation record – and especially the reports of panels that have to rely heavily on 'soft' inputs – is littered with complaints about the administration of the FPs. Late

[51] Barry Bozeman, 'Technology transfer and public policy: a review of research and theory', *Research Policy*, 29, 627-655, 2000
[52] FP4 Impact Austria
[53] FP4 Impact Growth
[54] FP4 Impact NNE

payments are also mentioned as a problem – one that is especially serious for SMEs. Complaints also abound about national level programme administration in national level evaluations. However, when beneficiaries are asked to compare the two they find the Commission's administration more burdensome. This difference seems to be driven by the greater scale and complexity of EU projects and the greater need for transparency in highly visible international projects. The only systematic benchmarking[55] of which we are aware focused on programme management proper, and found that the resources devoted to it were reasonable and the somewhat prolonged time scales could be explained by the complexity and international nature of the task. However, almost no work has been done to estimate the aggregate cost of FP Calls to proposers. FP4 Impact Germany suggests that the cost to applicants of writing proposals is equal to about 25% of the FP budget.

Proposal assessment seems to be well done. Internationally, the number of FP4 participations by each country was very strongly correlated with the number of researchers in the country, implying that a fair proposal assessment process was operating[56]. Proposal assessment normally receives good marks in the evaluations, either explicitly or implicitly, through judgements that the general quality and relevance of projects are good.

Understanding how the new instruments of FP6 were working was an interesting question for this study, but it was too early for there to be much evaluation feedback. However, while there is agreement that they are needed, the level of dissatisfaction with their implementation is high, with many – especially smaller – players seeing them as over-large and as placing them at a disadvantage. Costs of proposing have risen yet further and projects were artificially being inflated to meet the new requirements. While the new instruments were intended to increase flexibility, this was impeded by the Commission's financial and administrative rules. This is part of a more general issue the evaluations do not really get to: namely, the effect of these rules and procedures on frustrating new public management at the Commission. [57]

The FP5 Participation Finland study investigated how participants see FP6 against FP5. In Finland, excessive project size has not been considered problematic in FP5. One of the main changes in the shift from FP5 to FP6 is the requirement for larger consortia, which gives rise to several doubts among participants. The tendency towards larger project size is feared to

[55] James Stroyan and Erik Arnold, *Comparative Study of Administrative Burdens and Rules of Procedure between the EU Research Programmes and those of the Individual Member States*, Technopolis report to STOA, IV/98/06, STOA, European Parliament, 1999

[56] FP4 Impact Austria

[57] *Evaluation of the Effectiveness of the New Instruments of Framework VI*, Report of a High-Level Expert Panel chaired by Ramon Marimon, Brussels: EC, 2004; Technopolis, *Developing a Methodological Framework for High-Quality Assessment of the IST-RTD Effects (Results and Impacts) at the Strategic Objectives' Level*, Inception Report, Brighton: Technopolis, 2004

lead to problems in efficiency, division of labour, division of funding shares and co-ordination while increasing transaction costs in pursuit of improved efficiency, which, in the eyes of the participants, remains imaginary.

Two earlier evaluations suggest that larger networks such as those promoted by FP6 may be less effective than smaller networks. The FP4 Impact Biomed2 assessment states that scientific quality in the programme was negatively related to the number of partners in a project, and that the quality of the larger network projects was problematic. The FP5 Impact Growth study also found that larger networks (averaging 16 partners) generated limited impacts, compared with smaller ones.

Evaluations show that the FP was creating many aspects of European Added Value before that concept was widely discussed. As one example, **Exhibit 6** summarises the main points of added value reported by UK PMC members and UK programme participants.

As has been widely discussed in various panel reports, the concept of European Added Value (EAV) has been evolving and definitions have been inconsistent. Some of the evaluations in scope were launched, or relate to activities, prior to the widespread discussion of EAV or ERA.

Biomed2 projects had helped establish a European research culture, lasting collaboration and 25% of projects had set up shared facilities.[58] Finnish university researchers found that the FPs had strengthened their European networks during the 1990s, but this phase was now largely over.[59]

[58] FP4 Impact BioMed2
[59] FP5 Impact Finnish Universities

Exhibit 6 European Added Value (EAV) of Framework[60]

Augmentation of national funds
- Framework expands the funds available to national researchers over and above that which is available to them through national research funds alone
- Intellectual gearing
- Framework provides UK participants with access to foreign researchers and research outputs in a way that national funds cannot

Other types of added value
- Scale - pooling of resources as a means by which to increase investment in common European issues, from food safety to climate change
- Scope - pooling of competence as a means by which to increase the likelihood of a breakthrough in a given area from the economic manufacture of large structural composites to the sequencing of plant genomes

Strong added value in terms of the knowledge stock (science)
- Complex issues resolved more quickly and more thoroughly as a result of larger projects and portfolios multiple projects across successive Frameworks
- Status of knowledge accelerated through diversity and competition among national research traditions

Increasing added value in terms of support to EU policy
- EU regulates a growing number issues such as environmental protection or food safety
- Framework has made substantive additional investments in science, in areas such as climate change and infectious diseases
- Framework contributes to a more coherent EU view on risks and mitigation strategies
- However, arms-length involvement of policy makers limits real impact

Added value to EU businesses focused on key sectors
- Builds in-house competence, tools and de facto standards
- Strengthens international relationships

Danish participants believed the value added to the national research community by the FPs was improved access to international research.[61] Additionality seems also to vary according to national circumstances. In Ireland, where national funding for university research was very small before 2000, the additionality of FP projects was generally said to be very high (82% would not have gone ahead without FP money, according to FP4 Impact Ireland).[62]

In practice, many of the conventional benefits identified for international pre-competitive collaborative programmes imply EAV, in a European context: networking, especially international networking; facilities sharing; knowledge sharing; attaining bigger scale than is possible at the national level. These are often identified as benefits in project-level evaluations. We did not find any cases where it was argued EAV (or significant components of the concept) was **not** being achieved. Some of the evaluations argue that programmes and actions are promoting critical mass, or that they should do so in the future, but so far we have not seen any attempt at measurement of

[60] FP4-6 Impact UK
[61] FP4 Impact Denmark
[62] FP4 Impact Ireland

this within the evaluation sphere. (Of course, in other Commission studies, attempts are being made to understand critical mass through efforts such as the bibliometric 'mapping of excellence' approach.)

Evaluation surveys suggest that FPs tends to fund projects with high 'input additionality: that is, they would not have taken place without the EC funding, or if they had gone ahead this would have been in a more modest form: slower; smaller; with fewer partners, and so on. Economic theory[63] suggests there is a 'market failure' leading companies to under-invest in research of a comparatively long term or fundamental nature, as it is hard for them to monopolise the results. Correspondingly, evaluations surveys suggest that additionality is higher for more long-term parts of the FP than for those closer to market. It is also high for sectors that have low absorptive capacity and traditionally perform little R&D. Thus, in the construction industry, which has low technical and absorptive capacity, input additionality was high, with 75% of Growth projects being wholly dependent upon the subsidy for their execution.[64] BRITE-EURAM was a comparatively close to market programme with high SME involvement. Input additionality appeared much lower than in other parts of the FP. Only 45% of big companies and 51% of SMEs said they would not have gone ahead with the project in the absence of FP funding.

Exhibit 7 Additionality

	Participants that would have abandoned the project without funding		
	FP5	**FP4**	**FP3**
FP5 Impact	*58%*		
FP5 Impact Norway	*95%*		
FP4 Impact Denmark		*90%*	
FP4 Impact Austria		*70%*	
FP4 Impact Growth		*65%*	
FP5 Impact UK	*70%*		
FP4 Impact Ireland		*82%*	
FP5 Impact Finnish Universities	*75%*	*52%*	

RTD programmes like the FPs are often criticised for inadequate dissemination of results, and some such criticisms appear in the FP evaluations. An impact study of dissemination in the environment and sustainable development areas of FP5 found it was insufficiently targeted to specific audiences.[65] The NNE programme, too, was criticised for poor dissemination.[66] This was a frequent complaint that reappeared in other evaluations of programmes such as Growth, which tended to have more of

[63] Ken Arrow , 'Economic Welfare and the Allocation of Resources for Invention,' in Richard Nelson (Ed.) *The Rate and Direction of Inventive Activity*, Princeton University Press, 1962; see also Richard Nelson, 'The simple economics of basic scientific research,' *Journal of Political Economy*, 1959, vol 67, pp 297-306

[64] FP4 Impact Growth

[65] FP5 Impact of Dissemination in Environment

[66] FP4 Impact NNE

an innovation focus than a research focus.[67] In part this criticism arises from a misunderstanding, based on the popular 'linear model' idea that research produces results, which necessarily lead to development and commercial exploitation. The implication of this model is that results should be broadcast widely, so that they have a high probability of reaching those who can use them. However, the experience of RTD programmes over the past twenty or more years shows that this is a misunderstanding. Thus, it became a commonplace in the EU's SPRINT programme during the early-mid 1990s that involving end users to focus project goals on needs and focused rather than 'broadcast' dissemination were effective ways to reap implementation benefits from EU RTD programmes, but these lessons do not seem to have been well learnt at the level of the FPs. There is little point in such 'broadcasting' because most research results are irrelevant to most people most of the time. Some evaluations therefore stress the importance – especially in innovation-oriented SPs – of ensuring that end users and other 'problem owners' play a role in project definition, linking the production of research results directly to those able to use them. This underlines the importance in programme planning not only of consulting the potential research-performer beneficiaries but also of ensuring that those who will use results are also involved in specifying the kinds of results that will be useful.

5.3 Effects

The evaluation evidence about the effects of FP research is richer at the project level than at the FP level. There has been increasing effort to count the immediate outputs of FP programmes, for example in terms of patent applications, scientific journal articles, numbers of new participants not previously engaged in RTD activities. For example, NNE produced 400 patent applications, over 800 scientific journal articles, 1600 new actors not previously engaged in RTD activities. 30% of projects claimed to have made a technological breakthrough while 60% resulted in significant technical advances.[68] These kinds of indicators are useful as a 'pulse check': they show that the patient is alive and that some of the right kinds of processes are going on. However, in the absence of meaningful benchmarks or coherent ways of estimating socio-economic effects on the basis of these indicators they tell us little about overall performance, and are probably more useful for monitoring than evaluation purposes.

Because evaluations tend to be project focused, few explicitly consider the effects of the **portfolios** that are funded. In some cases, evaluators argue that lack of coherence in project portfolios reduces impact. The Growth evaluators point out[69] that, in the Production Technologies programme within Growth, there were many individual projects, which produced outcomes in areas such as health, assistance to disabled people, the environment and so on. However, the lack of coherence in the project

[67] FP4 Impact Growth
[68] FP4 Impact Growth
[69] FP4 Impact Growth

portfolio meant these effects would be scattered, largely invisible and not mutually supportive. The NNE impact study found the same, especially in relation to modelling studies.[70] Since few project portfolios are, in fact, well co-ordinated, this is likely to be a much more widespread problem than is apparent from the evaluations, and one that stems directly from Framework's character as a framework, rather than as a plan.

Beneficiaries can identify various benefits of project participation, and generally feel these outweigh associated costs. Benefits widely identified in evaluations are

- The R&D participants generally achieve their own goals, which tend to be knowledge and networking goals
- To a considerable degree, they also produce technological outputs, ranging from patents and articles to tools, techniques and methods which can be used in further research or in commercialisation
- The participants' speculation that socio-economic benefits will follow from their own technological work
- In a minority of cases, some evidence at the level of individual projects that such benefits are indeed produced, though the attribution of the more distant socio-economic benefits to the projects is problematic

There is a good degree of consistency among the benefits reported in different evaluations. It is not clear to what degree this is because the benefits are, indeed, similar or whether the similarity is partly a measurement artefact, since very similar questionnaires are used in almost all cases.

The FP3-4 Impact survey found the most important **expected** benefits (ranked) to be

- Improved international co-operation
- Enhanced career prospects for researchers
- Improved competitiveness of European industry
- Training and mobility of researchers

[70] F4 Impact NNE

Exhibit 8 shows that there is a good degree of consistency among the benefits reported in different evaluations.

Exhibit 8 Major benefits deriving from participation in FPs

	FP3-4 Impact	FP4 Impact Growth	FP4 Impact Ireland	FP5 Impact Norway
BENEFITS	- Knowledge and technological goals - Networking - Organisational and management goals	- Development of new tools and techniques - Enhanced skills of RTD personnel - Increase access to sources of expertise - Enhanced knowledge base	- Impact on scientific and technological standing - Capability-building goals	- Building up of competence and networks was successful - Scientific results of their participation was successful - The EU-participation laid the foundations for a new R&D project in the near future
	FP4-6 Impact UK	**FP4 Impact Austria**	**FP4 Impact Denmark**	
BENEFITS	- Improved the knowledge base of the participants, both in general terms and in relation to specific issues - Improve relationships and networks - Enhanced reputation	- Scientific reputation - Competitive position - Employment	- Development of new methods, - Scientific publications - Implementation of new technologies - Development of new products - Development of prototypes	

For Finnish universities, a factor analysis of responses concerning a long list of benefits showed the most important to be internationalisation, strengthened knowledge base, redirection of research (alerting the unit to new problems and providing the means to research into them), collaboration with end-users and commercialisation of research.[71]

The achievement of knowledge and networking goals almost universally reported in the evaluation record suggests that Framework works well in respect of its core functions – promoting research collaboration across Europe and strengthening S&T knowledge and capabilities. Taken together, the findings suggest that Framework is better at delivering the 'softer' knowledge and networking benefits sought most by the universities and public research institutes, and is less able to deliver the more concrete, commercially-oriented outputs and benefits desired by some parts of industry, especially SMEs.

Generally, participants say that the benefits of participation outweigh the costs, though the academic community is consistently more positive about this trade off than the industrial one.

[71] FP5 Impact Finnish Universities

Exhibit 9 Costs vs. benefits of FP Participation

	Benefits outweigh costs	Benefits equal costs	Costs outweigh benefits
FP5 Impact	55%	31%	14%
FP3-4 Impact	87%	4%	9%
FP4 Impact Austria	66%	17%	17%
FP5 Impact UK	69%	13%	17%
FP4 Impact Ireland	85% (est.)	10% (est.)	5% (est.)

The frequently-made claim[72] that the "most crucial test of programme effectiveness is whether or not participants reach their [own] project goals" is problematic. The international co-operation programme provides an extreme example of why this is the case, since the impact study[73] of the programme during FP4 found that EU participants' objectives were to do science, not to do development – which is the primary purpose of the programme. The justification for using taxpayers' money on R&D programmes is to reach socio-economic goals, not to redistribute income to programme beneficiaries.

This more important question of whether **programme** goals are reached is normally handled in very detailed project-by-project reviews at the level of individual action lines. At that level, goals tend to be to 'do research' in a particular area rather than to produce specific outputs or effects, and are hard not to satisfy. Adding up project impacts to draw conclusions about programme effects requires baseline and other contextual data, which are generally not considered in the evaluations, even if one or two evaluators are prepared to make a leap of faith and conclude that higher-level goals are satisfied. For example, the FP4 Impact Biomed2 assessment tackled goal fulfilment at the programme level by saying "scientific priorities in the various areas and sub-areas of the programme have in general been satisfactorily covered by the selected research projects."

[72] FP4 Impact Ireland
[73] FP4 Impact International RTD Co-operation

As shown in **Exhibit 10**, in all the cases considered here, knowledge-oriented goals were achieved.

Exhibit 10 Goal Attainment of FP Participants

	FP4-6 Impact UK	FP5 Participation Finland	FP4 Impact Ireland
GOAL ATTAINMENT	- Access research funding - Develop new or improved relationships or networks - Get answers to scientific or technical questions - Develop and extend internal knowledge and capabilities	- Networking - Monitoring scientific and technological development in the field	- Enhancing the knowledge base of participants
	FP4 Impact Austria	**FP4 Impact Growth**	
GOAL ATTAINMENT	- Gaining new knowledge - Expansion of the existing knowledge base	- Enhance existing knowledge base - Develop knowledge beyond state-of-art - Access complementary expertise - Enhanced skills of personnel - Explore new technology paths - Form new research networks	

Effects of FPs at national level are context-dependent. For example, they were big in Ireland because national R&D funding was so small, but over-reliance on EC funds also had the effect of fragmenting the Irish research community[74]. In the UK, the FPs were especially important because of the lack of national-level research and innovation funding aimed at companies. Naturally, the FPs' effects tend to be less obvious in large economies with well-established national R&D funding systems.

Socio-economic effects are the primary justification for the FPs, but are very hard to measure. The reasons for this are the methodological difficulties of doing so and, in many cases, the timing of the evaluation relative to the effects it would be desirable to measure.

A special study of the socio-economic dimension of community research in FP5 was able to trace effects of FP projects in European law via the Eurlex database. NNE demonstration projects tended to have short-term socio-economic effects, while the research projects took longer to pay off but were believed to have greater impact overall. In general, the energy area work reveals more clearly than other areas the importance of the context to obtaining socio-economic effects from technology projects. NNE was quite influential in the field of renewable energy, but was reported to have had little real-world impact on fossil-fuel energy use, because large and long-lived investments had locked the energy use system into particular

[74] Erik Arnold and Ben Thuriaux, *The Basic Research Grants Scheme: An Evaluation*, report by Technopolis to Forfás, Dublin: Forfás, 1998; James Stroyan, Ben Thuriaux, Erik Arnold, Alina Östling, Shaun Whitehouse, Sarah Teather, Kieron Flanaghan, Paul Cunningham, Mark Boden, Martin Visser and Anthony van Raan, *Baseline Assessment of the Public Research System in Ireland in the Areas of Biotechnology and Information and Communications Technology*, report to Forfás by Technopolis, PREST and CWTS, Dublin: Forfás, 2002

generations of technology. Nonetheless, half the demonstration projects and a third of the research projects in the NNE programme had led to cost and energy use reductions. Effects take time to materialise, so short term effects are more visible during and immediately after the projects, while it takes longer for any wider social effects to become visible.[75]

The international co-operation programme was said[76] to have a wide range of unquantifiable socio-economic benefits ranging from support of foreign scientists through a wide range of development objectives to effects on policy in poorer countries. It also gave the EU a positive image in third countries.

There is, however, clear evidence that the FP work can influence regulation and practice in new areas where norms are emerging, such as renewable energy or (in the past) new telecommunications standards, such as those emerging from the ACTs programmes FRAMES project and the UMTS multiple access air interface[77]. Only rather long-term retrospective study is likely to capture this kind of effect, or the effects the FP and its precursors had on the success of the GSM standard. It is harder to influence areas where there are large sunk investments or large, 'locked in' socio-technical systems.

An almost universal finding in the evaluations is that firms believe FP participation has improved their competitiveness. We read this to mean that they believe their **competitivity** (that is, their **ability** to compete, as opposed necessarily to their performance in competition) was affected, rather than their short term competitive performance. For example, in the FP-level survey[78] of FP3-4, 26% of industrial participants said their turnover increased as a result of their project; 60% reported increased 'competitiveness'. 'Innovation capability' rather than innovation was the main effect of FP participation by Norwegian companies.[79] In Ireland, the FP was credited[80] with allowing "Irish industry to raise technological capabilities on a broad front, with improved competitiveness a long-term rather than a short-tem consequence."

Some project-level evaluations use 'chain-link' logic to try to follow effects from projects to their outputs and their 'downstream' economic effects. This typically involves trying to assess the effects of projects on the cash flows of participants. In some cases, attempts are made to value other effects, such as estimating the value of networking induced by the projects. This can result however in a set of overlapping 'economic gains' in which benefits are

[75] FP4 Impact NNE
[76] FP4 Impact International RTD Co-operation
[77] Terttu Luukkonen, 'Technology and market orientation in company participation in the EU Framework Programme,' *Research Policy*, Vol 31, 2002, pp437-455
[78] FP3-4 Impact
[79] FP5 Impact Norway
[80] FP4 Impact Ireland

partly double-counted, and where the size of the benefits seems to depend not only on real socio-economic phenomena but also on the ingenuity of the evaluator in thinking up things to count. As a guide to the absolute size of effects, therefore, such approaches can be misleading[81].

However, when such methods are consistently used across different groups of beneficiaries, they do highlight the importance of 'leverage': the ratio of economic gains to inputs is higher for big companies (and for big countries) than for small ones. Unsurprisingly, the claimed effects of projects on the turnover of large companies are bigger in absolute terms than that on small companies.[82] The third BRITE-EURAM impact study found that while, on average, €1 invested by the EC in R&D support triggered €6.6 in "economic gain,"[83] the gain for large companies was €8.7. A later study associated lower economic returns with CRAFT projects than with others involving larger actors.[84] This finding sits awkwardly with the EC policy desire to integrate SMEs into the Framework Programmes. At the same time, it emphasises the importance of internationalisation of firms and the creation of a truly common market within Europe: the bigger the accessible market, the bigger the potential rewards to innovation.

Key FP effects were found in the way the research community behaves. Scientific and technological impacts were generally positive. Participants' behaviour was changed so that they would be more likely to work in networks than before. In some cases, the European research agenda was said to be changed. The FP-level survey of FP3-4[85] found 25% of respondents had achieved 'very positive', and a further 51% 'positive', scientific and technological impacts. Participants' behaviour was changed so that they would be more likely to work in networks than before. In FP4, 25% of Austrian participants[86] said that participation had had major effects on increasing their networking behaviour and 43% said there were minor effects. The evaluators credited[87] the NNE programme with "a significant contribution to developing EU leadership" in NNE research and strengthened academic-industry relations in 60% of the projects. The mid-term review[88] of the Ageing action in the FP5 Quality of Life programme credits it with changing the European research agenda in ageing. It argues (without quantification) that Europe does much less research in the area than the USA or Japan, and says that the action has raised the profile of ageing

[81] For a detailed critique, see Patries Boekholt, Maureen Lankhuizen, Erik Arnold, John Clark, Jari Kuusisto, Bas de Laat, Paul Simmonds, Susan Cozzens, Gordon Kingsley and Ron Johnston, *An international review of methods to measure relative effectiveness of technology policy instruments*, Report by Technopolis to the Ministry of Economic Affairs, The Hague, Min EZ, 2001

[82] FP4 Impact BRITE-EURAM

[83] This concept is problematic and tends to involve multiple counting of benefits. For a discussion, see Patries Boekholt et al.,2001

[84] FP4 Impact Growth

[85] FP3-4 Impact

[86] FP4 Impact Austria

[87] FP4 Impact NNE

[88] F5 Mid-term Quality of Life Action 6 Ageing

research in Europe so that more money will in future be devoted to it, and altered the character of ageing research by asking more holistic questions than before. As a result, European ageing research is becoming more interdisciplinary. FP was credited[89] in Ireland with playing "a vital role in maintaining and expanding the Irish research base". In the UK, it is seen[90] as important because there is so little national funding aimed at companies.

On the negative side, there are claims[91] that failure to co-ordinate genomic research at European level has left the European effort under-critical in this area. The Marimon report claims that the FP is inadequately co-ordinated with other sources of international R&D funding. These examples argue for stronger co-ordination in future.

When asked directly, participants tend not to feel the FPs have changed their research agendas. While they appreciated the FP as an additional source of funding, Danish university administrations said[92] that it had no significant effects on their research priorities. In Ireland[93] "there were some examples of academic institutions following FP agendas rather than their own, but these were the exception … there was little evidence that FP4 was distorting the country's scientific and technological base". The FP5 impact study on Finnish universities[94] says "respondents generally found that EU-funded research corresponds to the objectives of their units … less than 10% thought that EU participation has focused attention away from issues of national importance … few respondents thought that EU collaboration has brought some applied elements into their research. Rather, they considered the steering effect to be minor." Interpreting these claims is difficult. It is perhaps more likely that the FPs selectively attract and fund those whose research interests are in line with the foci of the Programme, rather than redirecting particular researchers from one research path onto another, so it is not clear that the effects of FP funding on the portfolio of projects would be visible to an individual project participant.

A key issue for the transition from FP4 to FP5 was whether the new FP was able to bring the R&D it funded closer to application than before. The 1999 Swedish study[95] on quality in the FP indicated that two thirds of researchers did the same sort of work nationally as in the FP, while one third did more applied work in the FP. The FP5 Finnish university impact study found that "respondents representing more application-oriented fields … were the most satisfied with EU programmes, whereas the respondents from more basic

[89] FP4 Impact Ireland
[90] FP4-6 Impact UK
[91] FP5 Impact Genomics
[92] FP4 Impact Denmark
[93] FP4 Impact reland
[94] Pirjo Niskanen, *Finnish Universities and the EU Framework Programme – Towards a New Phase*, VTT Technology Studies, Helsinki: VTT, 2001
[95] FP4 Scientific Quality of Swedish Participation

research-oriented fields were the most critical towards the relevance of EU programmes for their own fields."

Larger Finnish university groups were more likely to have FP projects than small ones. FP participants also had lower teaching loads, more research resources, more international visitors, more industrial collaborators and more TEKES funding than non-participants, suggesting that FP funds the more successful groups. The FP4 Impact Austria evaluation and the FP5 Impact Finland Universities study both found that bigger university groups were more likely to participate in the FP. This suggests that the FPs have for some time been reinforcing strength within the European research community, and to this extent building the foundations of a European Research Area.[96] The FP4 Impact Austria study also points out that large, R&D-capable companies are 'over-represented' among FP participants, suggesting that the more powerful R&D performers are being brought together by the FPs. According to a German industrialist quoted in the FP4 Impact Germany study[97], in the FP "the best in Europe find each other". However, looking at the project or programme level is insufficient to let us understand whether the increases in critical mass and research strength that appear to result from the FP are **sufficient** to make a difference to European scientific competitivity and performance. This would require a different kind of study.

The FPs' strengthening of more applied work and their focus on bigger and more successful R&D performers does tend to help structure the research community. Whether this happens enough to meet policy goals is an open question. (The 2000 Five Year Assessment suggests that the answer to this question is "No".)

Training in the broad sense of building experience is an implicit outcome found in almost every evaluation report. Thus, the FP-level survey of FP3-4[98] found that most respondents' believed the skills of their people had been increased by FP participation. A series of Finnish national impact studies has shown that FPs promote cross-sector collaboration and education and training of young scientists and engineers.[99] International co-operation promoted researcher mobility and built networks of Europe-friendly academics in the South, who formed network nodes for continued collaboration with EU researchers.[100]

One would expect the FPs to influence EC member state policy and affect private investment. However, no substantive evidence about these things was visible in the evaluation record.

---

[96] FP4 Impact BioMed2
[97] FP4 Impact Germany
[98] FP3-4 Impact
[99] Pirjo Niskanen, *Finnish Universities and the EU Framework Programme – Towards a New Phase*, VTT Technology Studies, Helsinki: VTT, 2001
[100] FP4 Impact International RTD Co-operation

Community social objectives (quality of life, health, employment, the environment) are rather distant from the individual projects, on which the evaluations tend to focus. Broad surveys find few employment effects, even if skills are improved. Impact studies of parts of Framework that focus on industrial and commercial results tend to identify some job creation, though the pattern is very skewed: most projects have little effect on employment; a very small minority have quite big effects. The impact study of the Telematics Applications Programme[101] indicated that the 100 projects considered had generated 137 jobs in one firm, 73 in a second and 40 in a third. The next 45 projects generated less than 30 jobs each and the bottom 34 projects produced less than 4 jobs each[102]. The FP5 Impact Growth study[103], which evaluated a large number of industrially oriented projects in some depth, claimed that the projects had generated 16 300 new jobs – but that 10 000 of them were in a single firm. (This skew is why the broad surveys do not pick up employment effects: where the unit of analysis is a participation in a project, the number of cases where significant impact can be claimed disappears into the noise.)

The role of SMEs in the FPs remains problematic. While there are corners of the FPs specially intended for them, in general SME participation is less successful than that of big companies and produces smaller impacts. There has been long-standing concern that SMEs are disadvantaged in the FPs and derive fewer benefits than big companies[104]. Involving SMEs and large firms in RTD programmes involves different kinds of risk. The economic fragility, and often the limited managerial capacities, of SMEs brings significant project risk (FP5 Impact Growth). Luukkonen argues that big companies are in the FPs to get knowledge, while SMEs are looking for markets[105]. In much of the FP, SMEs fare worse than larger companies. Technologically capable SMEs, such as academic spin-offs, appear to do better than more traditional small firms, but are still disadvantaged by their size and consequent instability. Arguably, this is a good reason why they should be funded – to reduce the risks of R&D – but it is also pointed out[106] that a failed FP project can bankrupt a small firm. The FP-level survey of FP3-4[107] found that 32% of those who found participation to have no or negative impacts were SMEs. Norwegian companies (mostly SMEs) were reluctant to participate in FPs because they produced intermediate knowledge outputs rather than results

[101] FP4 Impact Telematics

[102] In this study, the projects analysed were selected by the project officers in the TAP programme, so they are very unlikely to be representative of the programme as a whole. This tends to be confirmed by the widespread claims of job creation. Normally, in this kind of attempted impact assessment, less than half the beneficiary firms are able to quantify any of the benefits

[103] FP4 Impact Growth

[104] Philippe Laredo, The Networks Promoted by the Framework Programme and the Questions they Raise about its Formulation and Implementation, *Research Policy*, Vol 27, 1998, pp589-598

[105] Terttu Luukkonen, 'Technology and market orientation in company participation in the EU Framework Programme,' *Research Policy*, Vol 31, 2002, pp437-455

[106] FP4 Impact Denmark

[107] FP3-4 Impact

that could immediately be commercialised.[108] Outcomes for UK SMEs were poor.[109]

SME participants in the industrially-oriented BRITE programme tended to be R&D-intensive. Two thirds of them used project results directly in commercial applications.[110] The later Growth impact study[111] concluded that "without any doubt, the projects evaluated demonstrated that SMEs gained many benefits from their participation in the transnational research co-operation schemes." But SMEs were more likely to be frustrated in other, longer term oriented programmes. The FP4 Impact Growth study also tackles the CRAFT scheme, but in a very uncritical way. It points out that these projects tend to have a very pragmatic focus, are not especially high-tech and that two-thirds of participants are SMEs. It uncritically accepts that the international networking achieved by the scheme is a good thing *per se* and claims, without offering evidence, that CRAFT boosts R&D performance by SMEs. In contrast, the Norwegian FP5 impact study[112] questions the doctrine of SME participation: "there is no point to stimulate uncritically business, and especially SMEs, to participate in the EU's RTD activities. Businesses have to see clearly the relevance and the potential results from their participation. If not they do not participate."

The effects of the new instruments in FP6 make participation in many parts of the Framework less attractive to SMEs and tended to exclude new member state participation.[113]

The evaluations had little to say on other social policy concerns such as the new member states or on gender issues.

The Austrian FP4 Impact study[114] is unusual in tackling questions of gender, pointing out that only 8% of Austrian project leaders were women, while 35% of Austrian project participants were female.

A gender impact assessment of the Specific Programmes of FP5 was made in 2001. We have been able to obtain the study related to the IST (FP5 Gender Impact IST), which argues that – while gender was problematised in the Programme and targets set for women's participation – the 'gender-neutral' nature of IST implementation essentially reinforces the status quo. Gender issues are not systematically considered in the policy making process. The report proposes a number of measures that could help decrease the gender imbalances within IST.

[108] FP5 Impact Norway
[109] FP4-6 Impact UK
[110] FP4 Impact BRITE-EURAM
[111] FP4 Impact Growth
[112] FP5 Impact Norway
[113] Marimon Report
[114] FP4 Impact Austria

5.4 Feedback

The leap from the detailed level at which FP evaluations work to European or even national policy is a long one, which few evaluators attempt. Shifts in the nature of the Framework do not seem to have any basis in evaluation evidence – although clearly we would not expect evaluations to be the only source used. National FP impact studies tend to find there is a lack of national R&D strategy, and that such a strategy would be needed in order to make best use of the opportunities represented by the FPs.

If Member States want to know whether Framework is effective and efficient and whether it is worth investing more money in it then the Community needs much better evidence concerning what it does and what it achieves over and above the immediate outputs and impacts arising within project teams.

6 Conclusions

Our review leads us to two kinds of conclusions. Based on the evaluation materials, we can say something abut the way the FPs work and their impacts, as well as drawing conclusions about FP evaluation.

6.1 Conclusions about the Framework Programmes

The evaluation record suggests that the FP is a very useful and flexible device, allowing the EC to pursue a wide range of RTD-related programmes, but that it is flawed as a way to (plan to) reach specific policy goals. In practice, the low-level goals pursued in individual actions within SPs tend to be well-grounded with the stakeholder community, if not always with end users. Higher-level goals, however, are so abstract as to be untestable and hence remain untested. A key question is whether – if these higher-level goals were made more measurable – the **scale** of the FP would be adequate to its task. The question has not been rigorously addressed, but there is a lingering suspicion that the answer would be 'No'.

The quality of the science and technology undertaken in the FP appears broadly to be good. The FPs have consistently been dominated by the knowledge infrastructure of universities and research institutes, and this dominance appears to have become still more pronounced in FP6. There is little evidence based on experience yet about how well the new instruments of FP6 work. These focus on building critical mass by establishing larger RTD networks than were used in previous FPs. However, FP4 and FP5 experience was that large networks were less productive than small and medium-sized ones.

The Framework involves a strong 'core' of research institutions and companies that remains rather stable across FPs, spanning multiple projects and networks, and a rapidly churning 'periphery' of those who participate only once or twice. Significant parts of industry are not reached by the FPs.

Quite a number of large, R&D-performing European companies stay out. This is unfortunate in so far as large-firm participation provides high leverage for taxpayers' money. Since SMEs tend to be poorly served by most parts of the Programme, their limited participation is less problematic. However, it is worth noting that certain parts of the FPs are better oriented to SME needs than others, and that SMEs tend to have more favourable outcomes in these. If SME participation is an important objective of FPs, then these parts of the Programme need to be expanded. However, more study would be needed to determine whether SME participation in RTD is better handled as a policy issue at European, national or regional level.

FP project participants generally aim to produce 'intermediate outputs': that is, knowledge and networking outcomes that can subsequently be built upon, rather than results that can be commercialised in the short term. The projects have long generated the networking, resource sharing and mobile human capital that are now discussed as components of 'European Added Value'. FP programmes therefore tend to play a distinct role in the European Research and Innovation System, and do not simply duplicate national schemes. The FPs tend to 'structure' R&D within the ERIS by emphasising certain kinds of use-oriented research. There is relatively little evidence that they systematically change individuals' research agendas or priorities, as opposed to the 'style' in which they pursue these agendas (especially through increased networking).

While most participants primarily seek intermediate outputs, there is nonetheless a large volume of scientific and technological production. Innovation performance within any population tends to be skewed, and it is in the FPs, but it is clear that there are examples of significant wealth and job creation. Other impacts of the FPs include influence on regulation, the development and growth of certain kinds of new technologies, increased competitivity and skills, business and employment growth and influence over research agendas in certain areas.

Programme management and administration appear well done, but are widely seen as imposing considerable burdens on applicants. The formality and long-windedness of financial and legal procedures are among the reasons why, other things being equal, participants prefer national to EU funding. Participants therefore generally use FP funds for projects where international networking, resource sharing or other aspects of European Added Value are important to them.

It is clear that the FPs satisfy low-level process goals to 'do R&D' in various parts of science and technology and contribute to their high-level goals of strengthening the science and technology bases of European industry and contributing to the quality of life of European citizens. But it would be almost impossible to conduct large-scale R&D activities, guided by stakeholder consultation, without making such a contribution. Whether it does so in a cost-effective manner is not entirely clear but there is neither

evidence nor, on the face of it, any reason to believe that the EC does more or less well than national administrations funding similar programmes.

Key recommendations from the evaluation record include: the need for Member States to improve the way they develop national R&D funding strategies by taking better account of what is happening on the European level; and a need for FP planning to be better linked to evidence about problems and for increased stability within the FP itself.

6.2 Evaluation

There has been a rumble of criticism of the way in which Framework Programme (FP) evaluation is done over recent years, focusing on three sets of issues: methods; data availability and organisation; and the 'architecture' of the evaluation system itself.

Two recent studies EPUB[115] and ASIF[116] funded by the Commission pointed to a number of methodological deficits, which actually apply to all R&D evaluation – not just that of the FPs – and which are widely discussed elsewhere. They raise well-known issues such as attribution, the difficulty of assessing 'dead weight'[117], time scales, appropriate choice of methods and inadequate models of the relationship between R&D and other socio-economic variables. The authors of EPUB also rightly attack different evaluators for inventing their own definitions of impact and using them systematically to mislead policymakers.

ASIF goes into more detail about the difficulties of FP evaluation, and – like many of the FP Monitoring reports – refers to difficulties of document access, access to timely and adequately analysed data and independence from line management. It also argues that annual monitoring is too frequent to be useful.

The architectural issue is less discussed in print – though ASIF touches on it. The use of panels as the key arbiters in the evaluation process is an extreme extension of the principle of 'peer review'. The traditional, scientific use of experts is to make judgements about quality within the peers' areas of expertise, while this type of extension to the peer review principle means that most of the experts' work in the panel is to tackle issues outside their areas of expertise.[118] While the more extreme conflicts of interest

[115] Gustavo Fahrenkrog, Wolfgang Polt, Jaime Rojo, Alexander Tübke and Klaus Zinöcker (eds.) RTD Evaluation Toolbox: Assessing the Socio-Economic Impact of RTD Policies, STRATA Project HPV 1 CT 1999-00005, IPTS Technical Report Series, EUR 20382 EN, Seville: IPTS, 2002

[116] Luke Georghiou, John Rigby and Hugh Cameron (eds), *Assessing the Socio-economic Impacts of the Framework Programme (ASIF)*, University of Manchester: 2002

[117] Changes that happened at the same time as programme effects, but which were not actually caused by the programme

[118] For a useful discussion of this issue in the context of foresight, see Dennis Loveridge, "Experts and Foresight: Review and experience" PREST Discussion Paper Series paper 02-09, Manchester University: PREST 2002

highlighted by Georgiou[119] in 1995 have been addressed in the composition of recent panels, it is in practice difficult to assemble panels that exclude people from major beneficiary organisations. The panel approach is vulnerable to being exploited to promote disconnected political or policy agendas, as the ASIF study argues was the case with 5YA assessments prior to 2002. "Many of the recommendations drew not so much on an evaluation of past Framework activities but on the collective opinions and assessments of the Panel members concerning the general structure and organisation of science, technology and innovation in the EU."

The recent changes in the evaluation architecture, such as the increasing and more systematic use of strategic studies and the improved co-ordination of evaluation, address some of these issues. More fundamental evaluation problems, however, lie in the way the FPs' objectives are set, the disconnection between these objectives and the rapidly developing set of relevant policy objectives of recent years and the fact that the FPs obey what we have begun to term 'the law of the missing middle': that is, that the connection between the high-level goals and the low-level activities in the programme is not articulated.

The picture that this kind of meta-analysis of evaluations can provide quickly becomes outdated, and we are aware that improvements are in process. Nonetheless, the evidence from the past is of a fragmented pattern of evaluation, with some efforts to be experimental but a strong reliance on an approach that cannot readily address higher-level policy questions. Nor can the individual evaluations be 'added up' to provide a bigger picture. A common feature of many evaluations is the use of variants of a single questionnaire, rather than being a result of an architecture for evaluating the FPs. It is useful that FP-level surveys are complemented by national ones. However, the absence of studies from Southern and Eastern Europe means that the picture we have from this source is very biased towards Northern Europe.

As elsewhere, there is lots of mid-term and in-process evaluation; there is little evaluation done sufficiently far after the event that it is possible to understand impacts in a clear way[120]. The lack of specificity in goal formulation that arguably undermines the planning basis of the FPs also means it is difficult to evaluate the FPs' appropriateness, efficiency and effectiveness in reaching policy goals. Evaluation is forced down towards the level of projects, actions and specific programmes, focusing either on the quality and fulfilment of the technical objectives, on the one hand, or trying

[119] Luke Georghiou, 'Assessing the Framework Programmes – a meta-evaluation,' *Evaluation*, 1(2), 1995, pp171-188

[120] This type of evaluation, of course, has its own very significant set of methodological challenges and is by no means unproblematic – for example, because attributing causality is very difficult and because of information losses over time, so that the connections between projects and their impacts get lost. Some of the larger impact studies, for example in the Growth programme, have now begun to take a retrospective view at the project level

to get some measure of overall socioeconomic impact on the other, a process that involves huge and unresolved methodological problems and that ignores the interdependence of R&D and other activities.

But improvement is unlikely to be achieved without increasing the amount of evaluative effort devoted to answering policy questions that cut across the FP activities. More generally, the rush to get an up-to-the-minute understanding of impacts of activities that have not yet been completed leads to under-use of the evaluative effort that is already in place[121] "lack of exhaustive utilisation of data collected seems to be somewhat inherent in the overall organisation of EU's impact evaluations".

Policy makers and politicians naturally want to be able to show 'hard' and easily communicable evidence that their actions are effective, and few things look as hard as a number or are as easily communicable as 'for every Euro the Commission invests in R&D, the Community gets €x back'. Evaluators are therefore often driven to try to establish returns on investment for individual programmes or projects. This is problematic in at least four senses

- Normally, the effects need to be 'measured' before most of them have occurred
- Obtaining the necessary data and attributing causation are problematic and unreliable processes
- Innovation is a highly skewed process. Most innovation projects are financial flops, but the very successful ones tend to pay for themselves, for the flops and generate financial returns beyond this
- In real life, a lot of other things are needed over and above R&D results in order to generate economic benefits, such as investment, training, capital, and not least other complementary research results[122]

The use of single technique to evaluate a wide variety of programmes may be generating measurement artefacts: the programmes may look similar because we measure them all the same way, not because they actually are the same. The FP5 Growth Impact study is again suggestive in that the Standards, Measurement and Test activity is rather uncomfortably squeezed into the study's framework. These projects are done deliberately in order to build 'infrastructure' rather than to produce direct economic impacts. Properly understanding their socioeconomic effects would require a very different evaluation approach.

[121] ASIF Consortium, *Assessing the Socio-Economic Impact of the Framework Programme*, University of Manchester: PREST, 2002

[122] Using this kind of chain-link cause-effect thinking to work from projects to results to jobs has the perverse implication that, if results from another RTD programme are involved, it is unlikely that the evaluators will spot that these are already claimed by Growth and will instead credit a second programme with the same 10 000 jobs. At the limit, using such an approach to evaluation implies that the more evaluation you do, the bigger the apparent return on taxpayers' investment. That way madness lies.

Using the scientists and technologists who benefit from RTD subsidy programmes as the key informants about their effects is naturally problematic. In most EU cultures, based on cross-checking against other evaluations, the impression is that the information they provide is surprisingly honest. Perhaps we can credit the scientific tradition for this honesty. But we are left with serious questions about how much the project performers themselves can see, especially in relation to socio-economic impacts of their work. It seems likely that they will systematically over-attribute such effects to their own work; but also, systematically, be unaware of many if not most of these effects, especially as they tend to be indirect.

7 Where Do We Go from Here?

Our conclusions suggest that there are opportunities to improve both the way planning and evaluation are done in the FPs as well as the way these are organised. The FPs need better planning, with better links to an evidence base and an explicit intervention logic. Evaluation should be done within a systems perspective, and therefore should be linked with understanding and diagnosis of the ERIS. Individual evaluation studies should focus on limited numbers of tractable problems, rather than try to measure everything from first principles. New questions and methods will be needed and should be addressed within a clearer study and evaluation architecture in an organisation structure that links evaluation and FP planning. Increasingly, the Commission is taking a systems approach to the development of the ERIS, but this needs better to be reflected in the mechanisms through which the FPs are planned and through which their objectives are modified in line with changing policy needs. More transparent planning would have several important benefits

- Higher-quality, testable logics leading to an improved probability of reaching policy goals
- Evidence- and logic-based arguments to underpin the size of the budget needed by the FPs, shifting the balance of negotiation towards rationality and increasing the chances that the resources available for EU R&D policy are about the same size as the resources actually needed
- Improved evaluability, with corresponding benefits for improved processes, organisational learning and accountability to the taxpayer

If, as current innovation theory suggests, we live in a 'systems' world, then we need to have a systems view of causality. Neither FPs nor their components can **alone** cause the major changes in the European Research and Innovation System that are envisaged in the ERA, Lisbon and Barcelona agendas. (The same, of course, applies to other policies.) FPs may be necessary, but they are unlikely to be **sufficient**. For these reasons, too, more evaluative effort needs to be devoted to meso-level, policy-driven questions that are today not well- reflected in the goal structures of the FPs (Do Networks of Excellence create sufficient critical mass to build world class centres of excellence? Do the patterns of human mobility promoted by the

FPs improve the performance of the ERIS and / or increase cohesion within the EU?) We need to devote effort to understanding the things that the FPs can achieve directly, separately from the big socioeconomic impacts that they can only help achieve if other necessary conditions are satisfied.

This means we need to reduce our dependence upon linear, chain-link cause-effect analyses. Some of the things we believe we know about innovation suggest the same thing. Innovation is a complex, non-linear process, with emergent properties. We need therefore to understand (and to value) the intermediate outputs that almost all evaluations of pre-competitive collaborative R&D tell us are the main outputs. For example, we need to understand the 'Knowledge Value Collectives'[123] on which the FP projects operate and we need to know much more about human capital and mobility, since people are crucial as 'knowledge bearers'. We have to understand portfolio impacts, not to imagine that individual project impacts have much meaning. We have to evaluate over the long term as well as the short term. And we need to recognise that we do not have to evaluate everything. Intermittently quality-assuring processes that have been shown by evaluation to work is a good tactic for reducing the evaluation burden, allowing the resources freed up to be used to ask the more difficult evaluation questions.

Evaluation is, rightly, a sceptical profession, since its business is to test policies and interventions. This has two important consequences.

The scepticism means that little or nothing is taken for granted, and evaluators tend to try to demonstrate a complete chain of logic within their evaluations, connecting supposed causes and effects. It also means that – unlike science – evaluation is not cumulative. It does not establish theories and laws or use these to add new kinds of knowledge. Nor does it use *ceteris paribus* assumptions adequately. When we do scientific experiments, we create situations where we can manipulate a small number of variables and assume that the other variables and laws remain unchanged. Thus, we confidently take it for granted (with Newton) that if we apply a force to a body it will accelerate in a predictable way that is related to its mass. In R&D evaluation, we repeatedly observe relationships – for example, subsidising pre-competitive collaborative R&D increases human capital in the form of skills and skilled people – but cannot assume that this has consequences, since no evaluation of an R&D scheme is resourced simultaneously to develop theory about the connection between changes in human capital and overall socio-economic performance. Evaluations bog down in the need to prove everything from first principles.

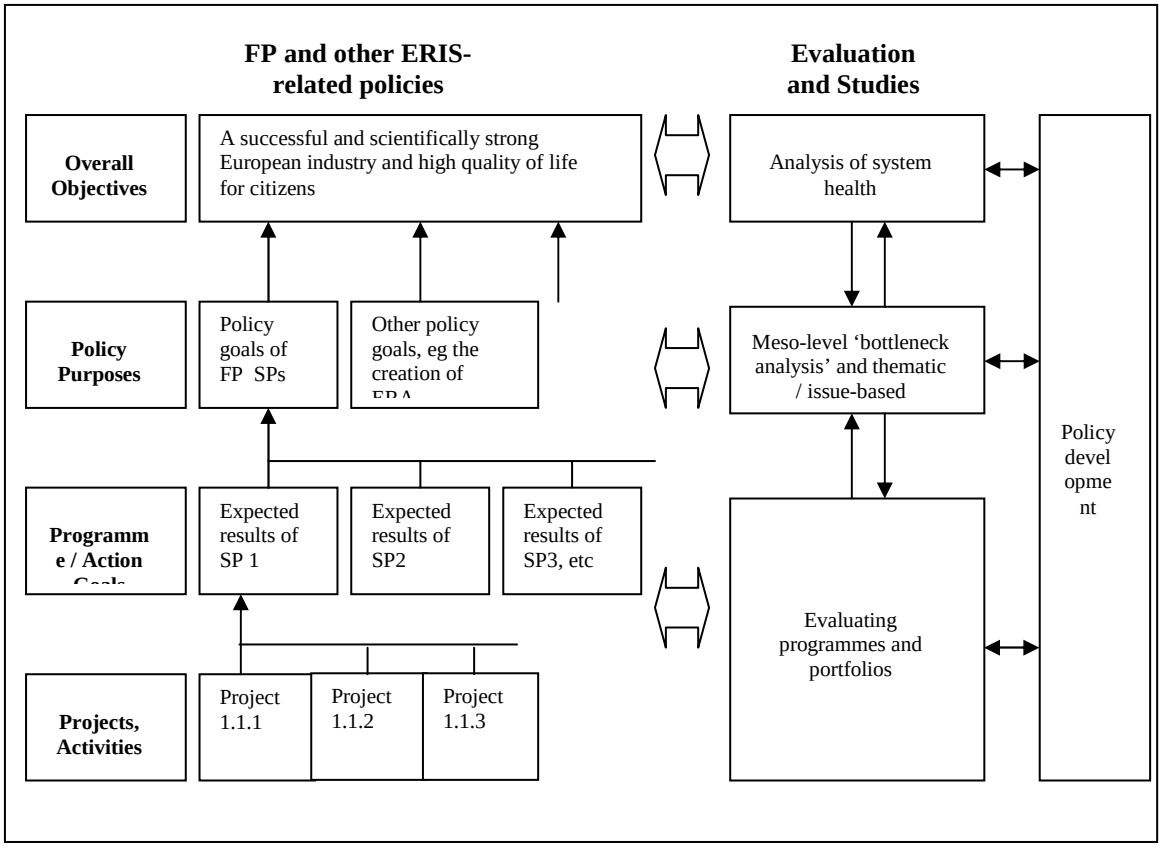
The professionalism is also a problem, because it involves the idea that evaluation is atheoretical and can therefore be done without reference to

[123] B. Bozeman, 'Technology transfer and public policy: a review of research and theory', *Research Policy*, 29, 627-655, 2000

theory about the domains in which it works. As a professional group, evaluators need to make this claim, so that they can move among domains. And the claim is not absurd. A great deal can be done using a combination of deductive logic and a set of social-scientific and statistical tools. But it follows, first, that professional evaluators often do not make use of helpful theory in the domains where they work, usually because they neither understand nor are aware of it, and, second, they are unaware of the theoretical assumptions they build into evaluations. One example of the latter is the use of patenting as an indicator of innovation, when those who work with patents know that patents are taken out not only to enable but also to prevent innovation, that most patents are never worked, that almost none repay their costs and that the role of patents varies among industries and between companies as well as over time. A second example is the repeated use of evaluation designs built on the so-called 'linear model' of innovation that was comprehensively rubbished in the research literature thirty years ago but that lives on in the popular and political imagination.

Given the massive complexity of the ERIS (and, for that matter, of the Framework Programmes), we will never be able to make the connections between low- and high-level activities and effects or between low- and high-level goals and policies without extending the repertoire of evaluation so that it can become cumulative as well as exploring from first principles. FP evaluation therefore needs better to be embedded in a systems perspective (**Exhibit 11**). Thus, evaluation at the level of projects, action lines and specific programmes needs to be complemented by evaluative studies that tackle the attainment of horizontal policy goals, such as critical mass, and which explore some of the many areas of R&D policy implementation (eg networks) which are still poorly understood and where better understanding would allow policy makers to improve the effectiveness of the programmes and instruments used.

Exhibit 11 Research and Innovation Policy Evaluation in a Systems World



There is a need for further review of the use of peer panels in the FPs. They have a crucial role in relation to scientific quality at the lower levels of the programmes. But the idea that a panel can, as the terms of reference of the 2004 Five Year Assessment panel demands, "peer review" the entire Framework Programme is a misapplication of the concept. No-one – no matter how clever – can be a 'peer' of something like the Framework Programme, any more than a lobster is a peer of a canary. The Assessment Panels may be important for **legitimation**, but they also need better systematic support, not least from a continuous programme of evaluation that links the programme and policy levels and is not solely focused on individual SPs or their component actions. The 2004 5YA moves in the right direction, but much more movement is needed.

The needed changes to planning and evaluation practice will not happen in a vacuum. Improvements in planning practice, especially in relation to performing ex ante analysis of the expected impacts of interventions and the adoption of a structure in which to plan such initiatives should provide one important basis. The Commission has been developing its evaluation practices since the late 1990s. From 2000, Directorates General of the EC were each required to have a properly staffed evaluation unit, operate an evaluation plan[124]. The Commission adopted[125] new Evaluation Standards

[124] SEC(2000)2203/5 of 13 December 2000

and Good Practice in 2002, and required their implementation by July 2003. Key requirements include

- Organisational separation of the evaluation function from 'line' departments
- Co-ordinating evaluation activities
- Anticipating decision-making requirements and ensuring that the evaluation plan aims to provide needed inputs to policy making
- Providing support in the formulation of policies and programmes and also in management processes
- Making adequate resource provision for evaluation activities
- Promoting standards and good quality in evaluation

There is now a clear need to devise a mechanism, which will allow more explicit planning of the FPs, using the emerging norms in the Commission for planning documents associated with interventions and requiring a higher level of specification of required outputs and outcomes. There must be a demonstrable intervention logic connecting actions, SPs and the Framework Programme as a whole and providing arguments to justify why particular amounts of resources should be devoted to particular parts of the FP. DG-Research's evaluation people need to be involved in this process, to support the development of testable goals and intervention logics and to ensure that the planning and evaluation parts of the policy cycle are interconnected.

In effect, this means interlinking the various policy initiatives operated by DG Research into a more coherent whole, with a common source of strategic intelligence. The DG already has significant strategic intelligence resources at its disposal, but a mechanism is needed to develop a programme of research and study that spans the collective needs of the different policies (FP, Lisbon, Barcelona, ERA and so on). Since 2003, a DG-wide evaluation plan, written by the evaluation unit, has been incorporated into the Annual Management Plan. Given an extended role for the evaluation unit, its capacity and workload should be reviewed with a view to increasing staff numbers. It needs to remain a 'staff' group and should operate in proximity to the FP planning function. However, it should remain organisationally separate, so that it does not in effect have to evaluate its own plans.

Evaluation at different levels and better integration with other kinds of strategic intelligence require new approaches. Of course methods that work should not be abandoned simply on the grounds of age, but there are many important policy questions that require different methods. Resources and freedom are needed here for experimentation, as well as routine execution of evaluations using established techniques. Better use should be made of the data that are collected, both by evaluators and in-house, so that more data can be re-used rather than re-collected.

[125] C(2002)5267

The range of new topics in FP evaluation that would be helpful is wide, and includes

- Country-level evaluations in Southern Europe and among new member states
- Improved mapping of R&D capabilities and potential (not just excellence) within Europe, and its benchmarking against global competition
- Modelling approaches to understanding impacts
- Horizontal, cross-FP evaluations of individual policy questions discussed above
- Longitudinal studies of how key institutions (R&D performers, companies) interact with the FP
- Macro and micro studies of RTD networks and networking behaviour
- Comparative studies about the effectiveness of instruments such as SME RTD supports at European, national and regional levels
- Long-term retrospective studies, to capture some major successes and explain the role of the FPs in them
- Mapping of human resources and Knowledge Value Collectives in relation to the FPs
- Detailed study of how actors interact with the FPs. In particular, exploring the way beneficiaries' Real Projects interact with the Framework and analysing incentives at the level of the individual participating group
- Analysis and benchmarking of administration and administrative costs
- Specific studies on gender, SME and new member state participation
- Conceptual exploration of the idea of scientific quality in the context of use-oriented R&D

A preparatory study is needed to identify how to improve FP planning and integrate it with FP evaluation. This should cover at least

- Process mapping of the evaluation and strategy planning processes affecting Framework
- Design of a logical hierarchy in which to define and manage evaluations and policy studies
- Implementation of the hierarchy as part of a wider planning framework

European Commission

EUR 21943— What the Evaluation Record tells us about the Framework Programme Performance

Luxembourg: Office for Official Publications of the European Communities

2005 — 47pp. — 21.0 x 29.7 cm

ISBN 92-79-00385-6

SALES AND SUBSCRIPTIONS

Publications for sale produced by the Office for Official Publications of the European Communities are available from our sales agents throughout the world.

How do I set about obtaining a publication?

Once you have obtained the list of sales agents, contact the sales agent of your choice and place your order.

How do I obtain the list of sales agents?

- Go to the Publications Office website http://publications.eu.int/
- Or apply for a paper copy by fax (352) 2929 42758

15

Ki-71-05-029-EN-C

Publications Office

Publications.eu.int

ISBN 92-79-00385-2

9 789279 003851