

IV

(Notices)

NOTICES FROM EUROPEAN UNION INSTITUTIONS AND BODIES

COURT OF AUDITORS

SPECIAL REPORT No 9/2007

concerning ‘Evaluating the EU Research and Technological Development (RTD) framework programmes — could the Commission’s approach be improved?’ together with the Commission’s replies

(pursuant to Article 248(4), second subparagraph, EC)

(2008/C 26/01)

TABLE OF CONTENTS

	Paragraphs	Page
EXECUTIVE SUMMARY	I-V	2
INTRODUCTION	1-14	3
The role of the European Union in supporting RTD activities	1-7	3
Evaluation and monitoring of FPs	8-14	5
AUDIT SCOPE AND APPROACH	15-18	7
AUDIT OBSERVATIONS	19-109	7
Intervention logic, objective setting and performance indicators	19-37	7
Evaluation strategy	38-49	10
Organising evaluation for the FPs	50-68	11
Evaluation methodologies and techniques	69-97	14
Disseminating and using evaluation findings	98-109	19
OVERALL CONCLUSION	110-118	21
<i>Annex I:</i> Measuring the impact and results of public intervention in RTD		23
<i>Annex II:</i> Provisions regarding monitoring and evaluation in legislative decisions for FP4 to FP7		25
<i>Annex III:</i> Main differences the in <i>ex-post</i> evaluation of research programmes between the EU (FP4 — FP6), USA and Canada		30
<i>Annex IV:</i> Possible panel structure for FP7		31
The Commission’s replies		32

EXECUTIVE SUMMARY

I. The Community RTD framework programmes (FPs) are the largest public source of research funding in the European Union. The Commission is responsible for managing the FPs and assessing their results.

II. The audit covered the monitoring and evaluation arrangements in place since 1995 for the last three programming periods (FP4, FP5, FP6) and also gives an outlook for FP7 (2007-2013). The audit was carried out with the EFTA Board of Auditors and the assistance of external experts.

III. The audit addressed the question of whether the Commission's approach to assessing the results of the FPs was adequate. In this context, the Court checked whether the Commission met the legal requirements for evaluation and ascertained whether its system for evaluation and monitoring met stakeholder expectations. In particular, the Court asked whether:

- (i) there was an underlying intervention logic to the FPs and their constituent programmes that could provide a sound basis for evaluation and monitoring;
- (ii) the Commission had defined an evaluation strategy for the FPs;
- (iii) the organisational structure of the Commission's evaluation system for the FPs ensured that evaluation was properly coordinated, adequately resourced and perceived by stakeholders as being sufficiently independent;
- (iv) methodologies used for evaluation were satisfactory; and
- (v) evaluation and monitoring activities provided information and analysis which could be used to improve ongoing programmes and prepare future FPs.

IV. The Court's findings and conclusions were as follows:

- (i) the lack of an explicit intervention logic and the presence of poorly defined programme objectives and weak performance measurement undermined effective monitoring and evaluation;
- (ii) the absence of a comprehensive evaluation strategy, agreed among the 'research DGs' implementing the FPs, resulted in inconsistent approaches between the different Commission services. In particular, this is illustrated by an under-evaluation of significant parts of the FPs;
- (iii) evaluation of the FPs was decentralised, the existing coordination mechanisms among the DGs implementing the FPs were not effective and the Commission's central services had no enforcement role. In general terms, however, the 'research DGs' have complied with the Commission's formal requirements on evaluation. The use of panels composed of high-level external experts appointed by the Commission for major evaluation exercises was well suited to the Community context. However, these temporary panels were not involved early enough in the process to influence evidence collection and had only limited possibilities to follow up relevant issues over time. By using only an overall FP panel (in addition to a panel for one specific programme), insufficient attention was given to specific issues related to the different scientific areas within the FPs;
- (iv) inadequate methodological guidance was provided, evaluators found difficulties in gathering relevant data and there were no evaluation studies that addressed the longer-term outcomes and impacts of the FPs, as opposed to short-term issues of programme implementation; and
- (v) as a result of the above, and because the timing of evaluations was often premature for addressing the most relevant issues, the Commission's evaluation system for the FPs was of limited usefulness to policy-makers, stakeholders or even the Commission itself. Ultimately, little is known about the achievement of programme objectives and the results of the FPs.

V. The Court recommends that:

- (i) intervention logic should be rendered explicit in future legislation. Underlying assumptions should be explained, the link between scientific and socio-economic objectives clarified and appropriate performance indicators developed;

- (ii) a comprehensive evaluation strategy should be developed by (and agreed among) the DGs implementing the FPs. In particular, this should entail a consistent approach with regard to the minimum level at which detailed evaluation must take place so as to take account of the specificities of each scientific field;
- (iii) consideration should be given to setting up a joint evaluation office for coordinating the 'research DGs' evaluation activities for the FP as a whole and creating a system of panels (and sub-panels) composed of external experts. These panels should be set up sufficiently early in order to provide effective assistance to the Commission, and continue to do so throughout the programming period and thereafter;
- (iv) the data requirements for evaluation and monitoring should be analysed properly and more extensive use should be made of other existing sources of data. The Commission should also rationalise the reporting requirements for participants. Finally, it should draw up a comprehensive evaluation manual for the FPs and develop a broader range of evaluation methodologies and techniques to be used for evaluations in this field; and
- (v) the Commission should establish the type and scope of evaluation that can be reasonably expected for the dates specified in the FP7 legislation and clarify how evaluations can be used to adapt programmes (learning programmes) and what contribution they can make to policy decisions.

INTRODUCTION

The role of the European Union in supporting RTD activities

1. Research and technological development (RTD) ⁽¹⁾ is the main driver of scientific and technological progress and innovation and, as such, is an important contributor to improvements in living standards. Public intervention in RTD is generally justified as a corrective to so-called 'market failures', with public support reducing uncertainty, providing investment where research is a public good or where there are significant externalities.

2. Article 163 of the EC Treaty establishes the Communities' role in supporting RTD, stipulating that 'the Community shall have the objective of strengthening the scientific and technological bases of Community industry and encouraging it to become

more competitive at international level, while promoting all the research activities deemed necessary ...'. Articles 164-173 determine the activities to be carried out in this respect and the scope and implementation of the multi-annual RTD framework programmes (FPs). The Commission is responsible for putting forward proposals for the FPs, whereas the FP decisions and the rules for participation and dissemination are adopted by co-decision of the European Parliament and the Council. The latter adopts the specific programmes alone. The FPs are implemented jointly by six Directorates-General (DGs), the so-called 'research DGs' ⁽²⁾.

3. Within the wide range of policies implemented in the European Union to strengthen innovation and competitiveness, the FPs are the most important financial instrument contributing to the Lisbon strategy ⁽³⁾ and the Barcelona objective ⁽⁴⁾ at the Community level.

⁽¹⁾ According to the OECD, research and development can be defined as '... creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications' (see OECD, Frascati Manual — Proposed standard practice for surveys on research and experimental development, Paris, 2002, p. 30). In the Community context, the acronym RTD is used, covering technology as well as research and development.

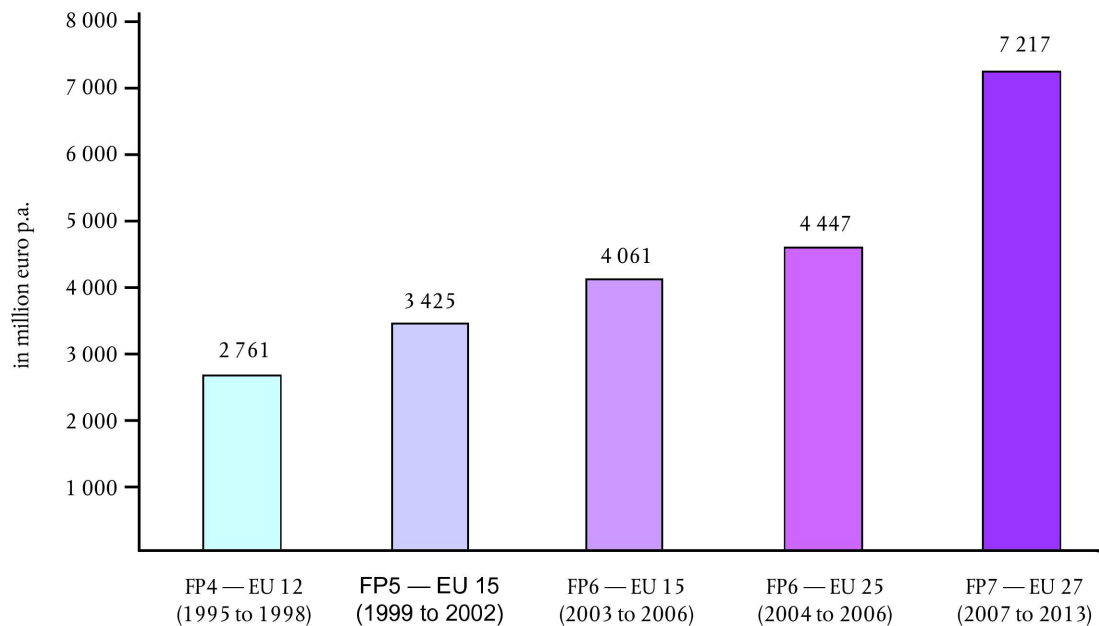
⁽²⁾ DG Research is the main Directorate-General managing and coordinating the FPs. Five other DGs also managing (parts of) specific programmes, the most important of which is DG Information Society and Media (the others being DG Energy and Transport, DG Enterprise and Industry, DG Fisheries and the Commission's Joint Research Centre).

⁽³⁾ In 2000, the Lisbon European Council set the goal of making the EU the most dynamic, competitive, sustainable knowledge-based economy by 2010, enjoying full employment and strengthened economic and social cohesion.

⁽⁴⁾ In 2002, the European Council agreed that overall spending on RTD and innovation in the European Union should be increased with the aim of approaching 3 % of GDP by 2010. It further clarified that two-thirds of this new investment should come from the private sector (see European Commission, 'Investing in research: an action plan for Europe'; Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions, COM(2003) 226 final, p. 7).

Figure 1

Average annual Community financial participation — FP4 to FP7 (EC)



Source: Legislative decisions for FP4 to FP7, see Annex II.

4. Through the FPs, the Community provides funding to researchers within the European Union, associated countries and international organisations. Their budgets have increased significantly over the years, reaching 7 217 million euro per year under FP7 as against 2 761 million euro under FP4 (see Figure 1). In terms of budgetary appropriations, they represent the largest area of direct centralised management within the Commission ⁽¹⁾.

5. The FPs differ from many national programmes in that they cover both basic and applied research with the participation of industry and public research organisations ⁽²⁾. They are made up of specific programmes (and sub-programmes) dealing with broad scientific fields. These programmes typically consist of a number of domains that concentrate on more specific scientific areas. These domains are then implemented through projects (called indirect actions) following calls for proposals. In this way, hundreds of individual projects are funded per domain.

⁽¹⁾ In the case of direct management, the Commission performs the operations required to carry out its budget implementation activities without any involvement by the Member States or third countries in which the recipients of the expenditure reside.

⁽²⁾ As of 2006, more than 22 000 public and private legal entities had participated in FP6 alone, generally within multi-partner consortia.

6. The overall share of the FPs in total public RTD funding within the European Union and its Member States ranges between 4 % and 5 %. Nevertheless, its impact on which kind of research activities are carried out within Europe is significantly larger. This is because Community grants generally provide only part of the total funding for a project. In addition, when applying for funding, researchers need to demonstrate that their projects address the objectives specified in the calls for proposals. In this way, the FPs set incentives to orient RTD activities towards specific EU objectives.

7. In recent years significant changes have been made to the FPs, due to:

- new political challenges (the Lisbon strategy and Barcelona objectives, the establishment of a 'European Research Area' (ERA) and the increasing use of the open method of coordination ⁽³⁾ in research and innovation policy),
- the changing orientation of FPs, from simply solving technical problems under FP4 and addressing specific social problems under FP5, to encouraging the co-operation of research organisations within the ERA and contributing to the establishment of a knowledge-based economy and society under FP6 and FP7,

⁽³⁾ This is a method of cooperation based on the principle of voluntary intergovernmental cooperation. It rests in particular on mutual learning, benchmarking and the sharing of best practice.

- longer programming periods (previously 4 years, now 7 years for FP7 in line with the 2007-2013 Financial Perspectives),
- more Member States and participating countries,
- increased funding, in particular under the current Financial Perspectives,
- the setting up of the European Research Council (ERC) and the delegation of programme management activities to executive agencies starting under FP7, and
- the increasing use of large-scale funding instruments (e.g. Networks of Excellence, Integrated Projects, Article 169 projects, Joint Technology Initiatives).

Evaluation and monitoring of FPs

8. The Commission has responsibility for programme implementation and the assessment of programme results ⁽¹⁾. Within the Commission, evaluation is decentralised to the DGs, with the central services (in particular DG Budget and the Secretariat-General) providing support and coordination.

9. This report employs the Commission's own definitions of 'evaluation' as a 'judgement of interventions according to their results, impacts and needs they aim to satisfy' ⁽²⁾, and 'monitoring' as providing information on the implementation of ongoing programmes and projects to address management needs ⁽³⁾. This is also how the two activities are distinguished in the current FP7 legislation.

⁽¹⁾ Council Regulation (EC, Euratom) No 1605/2002 of 25 June 2002 (OJ L 248, 16.9.2002, p. 1), Articles 27, 48, 59 and 60.

⁽²⁾ European Commission, DG Budget, Communication to the Commission from Ms. Grybauskaitė in agreement with the President, 'Responding to strategic needs: reinforcing the use of evaluation', SEC(2007) 213, 21 February 2007; European Commission, DG Budget, Communication to the Commission from Mrs Schreyer, 'Focus on Results: Strengthening Evaluation of Commission Activities', SEC(2000) 1051, 26 July 2000.

⁽³⁾ Decision No 1982/2006/EC of the European Parliament and of the Council of 18 December 2006 concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013) (OJ L 412, 30.12.2006, p. 1): Article 7(1); the monitoring information to be provided by the Commission to Programme Committees is defined in the specific programme decisions (see for example Council Decision 2006/971/EC of 19 December 2006 concerning the Specific Programme Cooperation implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013) (OJ L 400, 30.12.2006, p. 86): Article 8(4) and Annex V).

10. Hereinafter, the following terminology is used to describe the results of RTD activities:

- 'outputs' are knowledge, goods and services directly produced as a consequence of the public intervention (publications, conference papers, patents, licences, prototypes, standards, trained scientists, etc.),
- 'outcomes' are the delayed result of the users' processing of outputs, and
- 'impacts' are the long-term socio-economic and political changes the public intervention brings about.

Box 1

Example of the 'Global System for Mobile communications' (GSM)

In this example, the 'outputs' of Community funding were the immediate project results, such as the development of the GSM standard. The 'outcome' was its acceptance as the digital cellular telephony standard within Europe and beyond, which contributed greatly towards the development of the telecommunication industry. Finally, 'impacts' were the economic, social and political consequences of the international use of this technology. This includes contributing towards job creation within the telecommunications and related industries, and the growth of the importance of this industrial sector for the European economy. GSM has also ultimately provided a new way for people to communicate with each other.

11. Assessing those results, however, poses a number of methodological difficulties (see *Annex I*). Socio-economic changes are complex and, more specifically, the relationship between a research activity, and the outcomes and impacts that are observed is often difficult to trace. While representing a considerable challenge, these methodological difficulties do however not necessarily preclude effective evaluation.

12. The Commission has had a system to monitor and evaluate the FPs in place since the 1980s ⁽⁴⁾. Starting with FP4 in the mid-1990s, the monitoring and evaluation system consisted of a number of partly linked activities, including:

- annual monitoring (mainly of programme implementation), and
- five-year assessments (5YAs) of FP activities (and, as from FP6, additional assessments of specific issues).

Both activities had to be carried out with the assistance of external experts (see *Figure 2* and *Annex II*).

⁽⁴⁾ European Commission, Communication to the Council concerning a Community Plan of Action relating to the Evaluation of Community R&D Activities for the Years 1987 to 1991, COM(86) 660 final, 20 November 1986, in response to the Council Resolution of 28 June 1983.

Figure 2

FP4 to FP7 — key requirements for monitoring and evaluation

FP4 (1995-1998)	FP5 (1999-2002)	FP6 (2003-2006)	FP7 (2007-2013)
two evaluation exercises (mid-term and final) , in each case for its 20 sub-programmes plus the FP; final evaluation overlaps with mid-term of FP5	two simultaneous evaluation exercises (mid-term) , in each case for its 7 Specific Programmes plus the FP	one evaluation exercise for the FP and its 3 Specific Programmes (mid-term), <i>but not specifically for all 20 sub-programmes</i> , complemented by specific evaluation and regular progress (i.e. monitoring) reports	at least five evaluation exercises (ex-ante, mid-term and final) including specific progress report before interim evaluation
<p>1. Ex-ante: no</p> <p>2. Mid-term: — 5YA: mid-way through FP, prior to subsequent FP (1997)</p> <p>3. Final: — on completion on FP (1999)</p>	<p>1. Ex-ante: no</p> <p>2. Mid-term: — 5YA: prior to presenting proposal for subsequent FP (2000) — progress review mid-way through FP (2001)</p> <p>3. Final: no</p>	<p>1. Ex-ante: no</p> <p>2. Mid-term: — evaluation of effectiveness of instruments (2004) — 5YA, prior to presenting proposal for subsequent FP (2004) — regular progress (i.e. monitoring) reports</p> <p>3. Final: no (<i>modified by to FP7 Decision</i>)</p>	<p>1. Ex-ante: incorporated in 'Impact Assessment of the Commission's proposal for FP7' (2005)</p> <p>2. Mid-term: — <i>ex-post</i>, two years after the end of previous FP (2008) — progress report before interim evaluation (before 2010) — interim evaluation (2010)</p> <p>3. Final: two years after completion of FP (2015), supported by specific studies, the interim evaluation and other evaluation activities carried out throughout the FP period</p>

Source: see Annex II.

13. To date, three 5YAs have been carried out ⁽¹⁾, leading to the 'Davignon', 'Majó' and 'Ormala' reports (1997, 2000 and 2004 respectively). In addition, DG Information Society carried out targeted assessments for its specific programme. It should be noted that 5YAs were discontinued under FP7 legislation, thereby removing the constraint that major evaluation exercises had to concentrate on the preceding five years and, as a result, had to cover parts of two successive programming periods. In 2004, the 'Marimón' report ⁽²⁾, a specific evaluation regarding the effectiveness of the new instruments introduced under FP6, was also published.

14. The Commission's framework for evaluation activities was reviewed in 2000 and 2002 ⁽³⁾. These communications were updated in 2007 taking account of the consequences of the Commission's administrative reform and the new Financial Regulation ⁽⁴⁾. In these documents, the role and function of evaluation is described in the context of the Commission's overall planning, programming and reporting cycle and a set of 21 evaluation standards and 18 models of 'good practice' are specified. Furthermore, these documents discuss the use of evaluation in the preparation and simplification of EU legislation.

⁽¹⁾ European Commission, 'Five-year Assessment of the European Union Research and Technological Development Programmes, 1992-1996'. Report of the Independent Expert Panel chaired by Viscount Davignon (1997); European Commission, 'Five-year Assessment of the European Union Research and Technological Development Programmes, 1995-1999'; Report of the Independent Expert Panel chaired by Joan Majó (July 2000); European Commission, 'Five-year Assessment of the European Union Research Framework Programmes, 1999-2003'; Report of the Independent Expert Panel chaired by Erkki Ormala (15 December 2004).

⁽²⁾ European Commission, 'Evaluation of the Effectiveness of the New Instruments of Framework Programme VI, 21 June 2004, Report of a High-level Expert Panel chaired by Professor Ramon Marimón.

⁽³⁾ European Commission, DG Budget, Communication to the Commission from Mrs Schreyer, 'Focus on Results: Strengthening Evaluation of Commission Activities', July 2000, SEC(2000) 1051; European Commission, DG Budget, Communication for the Commission from the President and Mrs Schreyer, 'Evaluation Standards and Good Practice', C(2002) 5267 23.12.2002.

⁽⁴⁾ European Commission, DG Budget, Communication to the Commission from Ms. Grybauskaitė in agreement with the President, 'Responding to strategic needs: reinforcing the use of evaluation', SEC(2007) 213, 21 February 2007.

AUDIT SCOPE AND APPROACH

15. The main objective of the audit was to determine whether the Commission had an adequate approach to assessing the results of the RTD framework programmes.

16. To this end, the Court assessed the extent to which:

- an explicit intervention logic underlies the FPs and their constituent programmes around which outputs, outcomes and impacts can be measured and assessed, thereby providing a sound basis for evaluation and monitoring,
- an evaluation strategy has been defined for the FPs promoting effective and consistent evaluation,
- the organisational structure of the evaluation system ensures that it is appropriately coordinated across the FP, adequately resourced and also perceived by stakeholders as being sufficiently independent,
- the methodologies used for evaluation ensure that relevant issues are addressed in an appropriate way, so that evaluation is perceived as credible, and
- evaluations and monitoring activities provide information and analysis which can be used to improve ongoing programmes (learning programmes) and prepare future FPs.

17. This report covers the relevant arrangements for FP4 (1995-1998), FP5 (1999-2002) and FP6 (2003-2006). It also discusses the legal requirements currently in force under FP7 (2007-2013), but does not provide an assessment of the Commission's FP7 'Impact assessment and *ex-ante* evaluation' ⁽¹⁾. Its focus is on the monitoring and *ex-post* evaluation of programmes, and not the *ex-ante* evaluation of project proposals or *ex-post* evaluation of the projects themselves.

18. The audit was carried out with the participation of the EFTA Board of Auditors and with the assistance of external experts. Audit evidence was collected through:

- audit work at the Commission (documentation review and staff interviews), involving the identification of legal and statutory requirements, a review of objectives and performance criteria, the analysis of more than 80 monitoring and evaluation reports and an analysis of the organisation of evaluation activities within the Commission,
- interviews with more than 90 experts identified as stakeholders in Member States and associated countries (Czech Republic, Germany, Estonia, Spain, France, Italy, Luxembourg, Netherlands, Austria, Finland, Sweden, United Kingdom and Norway),

⁽¹⁾ European Commission, Staff Working Paper, 'Impact assessment and *ex-ante* evaluation', SEC(2005) 430.

- surveys with institutional stakeholders at the European Parliament's ITRE ⁽²⁾ and the Council's CREST ⁽³⁾ committees,
- a consultation with more than 65 experts in the field of research evaluation in the USA and Canada as an international comparison group,
- a review of academic literature on current trends in research evaluation, and
- the discussion of the audit findings, conclusions and recommendations with research evaluation experts from Europe and North America during a series of facilitated focus group meetings in February 2007.

AUDIT OBSERVATIONS

Intervention logic, objective setting and performance indicators

19. Effective evaluation and monitoring requires clear objectives and relevant performance indicators at policy, programme and project level against which the results of a public intervention can be assessed. Good practice in other EU budgetary areas and further afield suggests that these should be defined within the framework of an intervention logic.

20. Intervention logic can be defined as a set of hypothetical cause and effect linkages that describe how an intervention is expected to attain its global objectives ⁽⁴⁾. The Commission's guidelines for evaluation state that intervention logic provides 'the conceptual link from an intervention's inputs to the production of its outputs and, subsequently, to its impacts on society in terms of results and outcomes' ⁽⁵⁾. This report uses this definition throughout.

⁽²⁾ The European Parliament's 'Industry, Research and Energy' (ITRE) Committee is (in particular) responsible for industrial policy, research policy (including the dissemination and exploitation of research findings), space policy, the activities of the Joint Research Centre, the Euratom Treaty, and the information society and information technology.

⁽³⁾ The CREST ('Comité de la Recherche Scientifique et Technique' — Committee for Scientific and Technical Research) was set up by a Council Resolution dated 14 January 1974 and is made of representatives from national authorities who are responsible for the scientific and technological policies of Member States. CREST has an advisory role, informing the Commission and the Council, in particular on subjects relating to the coordination of national R&D policies and the monitoring and evaluation of the FPs.

⁽⁴⁾ Logic models showing how inputs in the different activities lead to expected results are the most appropriate way of illustrating intervention logic. These models are capable of explaining in a clear manner the assumptions underlying the way in which a programme is designed and how overall objectives are broken down into operational objectives.

⁽⁵⁾ European Commission, DG Budget, 'Evaluating EU activities: a practical guide for the Commission services', July 2004, pp. 87 and 106.

21. The Court assessed the extent to which intervention logic was established for the FPs and its constituent programmes, thereby providing a sound basis for evaluation and monitoring.

Intervention logic for the FPs has not been explicit

22. Defining intervention logic should be a key element of programme design from the outset. Article 163 of the Treaty, stating the need to strengthen the scientific and technological base and become more competitive, should provide the starting point for such an intervention logic for the FPs (see paragraph 2). This would need to be made more specific with regard to the different scientific areas covered by the FPs and the different instruments used.

23. However, an explicit logic model explaining how the FPs and their constituent programmes contribute towards scientific and technological progress and innovation is lacking at all levels.

24. The absence of intervention logic for the FPs was also highlighted by the 2004 5YA panel, which stated that '... panels like ours are asked to fill a gap between, on the one side, evidence mainly collected at project level and, on the other side, the higher level socio-economic goals of research policy. However, at the moment the link is difficult to make due to the way the FP is planned. It lacks an explicit logic connecting the highest objectives to the specific research and knowledge goals' ⁽¹⁾.

25. A different approach is taken in some Member States (e.g. The Netherlands, Finland) and the USA (see Figure 3), where defining intervention logic is much more commonplace and in Canada where it is mandatory for all programmes (see Annex III). Indeed the use of intervention logic in programme design there results in more focused and better structured programmes. It is also being introduced by the Commission in other Community policy areas (e.g. external aid) ⁽²⁾.

Figure 3

Definition of a hierarchy of objectives based on logic models — example of US agricultural research

Policy	Programme	Project
Policy objective: to contribute to the improved livelihood of farming families		
Purpose: increased agricultural production, productivity and incomes among farming households	Programme objective: to contribute to increasing agricultural production, productivity and incomes among farming households	
Expected result: the use of improved agricultural technologies increased among targeted farmers	Purpose: increased use of improved agricultural technologies by farmers (e.g. rice)	Project objective: to contribute to increased use of recommended improved technologies
	Expected result: recommendations for targeted farmers provided/disseminated	Purpose: recommendations provided for improved technologies suitable for targeted farmers
		Expected results (e.g.): 1. technologies identified based on farmers priorities 2. technologies generated and adapted 3. technologies verified in farmers fields

Source: US National Agricultural Research Council.

⁽¹⁾ 'Five-year Assessment of the European Union Research Framework Programmes, 1999-2003', Report of the Independent Expert Panel chaired by Erkki Ormala (15 December 2004): Section 6 'Evaluating the Framework Programme', p. 19.

⁽²⁾ Recipient countries of external aid are required to present logic models for their Community funding (European Commission, Aid delivery methods, Volume 1, Project cycle management guidelines, March 2004, Section 5: The logical framework approach).

26. Not codifying the underlying rationale for public intervention and the assumed causality links (according to which the programme is then planned in detail, executed and assessed) undermines the effective evaluation and monitoring of the FPs. This is particularly due to how programme objectives and performance indicators are defined.

Without an explicit logic model it is more difficult to set coherent programme objectives

27. Not making the underlying rationale of the public intervention explicit makes it more difficult to set and structure programme objectives coherently. This is demonstrated by the fact that:

- it is not clear how the FPs' programme objectives relate to policy objectives, such as the overall objective in Article 163 of the Treaty, those stated in the Lisbon strategy and the Barcelona declaration (see paragraph 2 and 7),
- the relationship between individual programme objectives, both at the different layers of the legislation (such as FP decision, specific programme decision and rules of participation and dissemination) and between the basic acts and the Commission's work programmes, is not always apparent.

28. The Commission has attempted to set out operational objectives in the work programmes based on the programme objectives defined in the FP legislation. However, in almost all cases, these initiatives have failed to clarify the programmes' underlying logic and the objectives stated in the work programmes were not specific, measurable, achievable, relevant and timed (SMART). In 2003, this became a legal requirement under the Financial Regulation (1).

29. The extent to which policy objectives can be traced through the specific programmes, work programmes and calls for proposals to the individual projects co-financed by the Community is unclear. Unless these links are clarified, the contribution of the FPs and their constituent programmes to the achievement of these policy objectives cannot be assessed. The audit found that little or no convincing evidence about the achievement of longer term objectives, both at project and programme level, has been provided by the Commission in its evaluations (see paragraph 82).

The absence of an explicit intervention logic and clear objectives make it more difficult to develop performance indicators

30. The absence of an explicit logic model and clearly spelled-out programme objectives has also impeded the development of performance indicators for monitoring the achievement of programme objectives in terms of outputs, outcomes and impacts, to the extent that, to date, such 'effectiveness' indicators have not been defined at all for the FPs (see paragraph 73).

(1) Regulation (EC, Euratom) No 1605/2002, Article 27(3).

31. In this situation, the review of ongoing projects funded by these programmes and the *ex-post* evaluation of completed ones is also rendered more difficult and contributes to the difficulties in aggregating evaluations from the project to the programme level (see *Annex I* and paragraph 84).

ABB/ABM as currently implemented is an unsuitable tool for performance measurement of the FPs

32. Some indicators that track the efficiency of programme implementation (e.g. time-to-contract, time-to-payment) are found in the Commission's annual monitoring reports and also in the context of the Commission's Activity-Based Management (ABM) system.

33. However, up to now, the 'research DGs', even when implementing the same specific programme (or sub-programme), have not applied a standardised or comparable Activity-Based Budgeting (ABB) nomenclature and related indicators. Therefore, these indicators cannot be used to compare performance between and within programmes.

Box 2

ABB/ABM as currently implemented by the 'research DGs'

in the ABB/ABM system, the FPs are covered by two different policy areas within which the FP activities are presented according to a different DG-specific nomenclature (2). This even applies to those specific programmes and sub-programmes implemented by more than one DG. ABM objectives were defined according to different underlying principles (3) and, as indicated by the Commission's Internal audit service (4), were mostly formulated in general terms, describing aspirations or activities. Indicators were generally input-oriented (e.g. number of meetings, etc.). Some regarded process efficiency (e.g. time-to contract, time-to-payment), but were not comparable between programmes and from one DG to the other due to the differences in the ABB nomenclature, in DG-specific assumptions (such as the beginning and end of specific administrative processes) and in the underlying population (e.g. type of grant agreement, type of payment) that was taken into consideration for their computation.

(2) In 2006, policy area '08 research' was structured on the basis of the specific programmes (and sub-programmes) of the current FP6 programming period (and as a consequence replaced by a different structure in 2007), whereas within policy area '09 information society' the current and previous specific programme were put together as one ABB activity (European Commission, DG Research, Directorate A: Coordination of Community activities, Planning, programming, evaluation, Annual Management Plan 2006, 20.12.2005; European Commission, DG Information Society and Media, 2006 Annual management plan, 23.12.2005).

(3) In the case of DG Research, ABM objectives (and the indicators used to measure their achievement) changed from one year to the next. DG Information Society, on the other hand, used multi-annual objectives, which however, were not the same as the programme objectives (or those specified at work programme level).

(4) European Commission, Internal audit service, Review of effectiveness and efficiency of the SPP/ABM cycle — SG & DG Budget, 2005: paragraph 4.1, p. 31.

Recommendations

34. In future legislative proposals to the Council and the European Parliament, the Commission should set out an explicit intervention logic for the FP as a whole and its constituent parts.

35. For FP7, the Commission should:

- clearly describe the objectives of the FP and its constituent specific programmes (and sub-programmes) and how they should be achieved,
- clarify the link between the scientific and socio-economic objectives for each specific programme (and sub-programme), and
- define relevant operational objectives.

36. For its FP7 monitoring system, the Commission should put in place a limited, but balanced, set of performance indicators to measure the efficiency and effectiveness of programme implementation for each specific programme (and sub-programme).

37. The 'research DGs' should harmonise their ABM/ABB nomenclature to make performance indicators comparable both over time and, where possible, between specific programmes (and sub-programmes).

Evaluation strategy

38. An evaluation strategy provides the conceptual framework within which evaluation activities are designed, planned, executed and used. Good practice in other EU budgetary areas and further afield suggests that such a strategy should consider the main legal, organisational and methodological issues surrounding programme evaluation. This includes what evaluations are to be carried out, by whom and when, how data are to be collected, what methodological approaches are to be used and how findings are to be communicated and followed up.

39. Having such a strategy in place, and making sure that it is implemented, is a pre-requisite for embedding evaluation in the Commission's organisational culture and decision-making process. The Court recognises, however, that given the diversity of the scientific areas covered by the FPs, such a strategy should provide an appropriate degree of flexibility to each of the DGs responsible and should not be overly prescriptive.

40. The Court checked the extent to which the Commission had in place such an evaluation strategy for the FPs.

Absence of a comprehensive evaluation strategy for the FPs

41. For the FPs, the Commission first made public its approach to evaluation in the 1980s ⁽¹⁾. This was updated in 1996, when the Commission informed the European Parliament and the Council of what it then regarded as the relevant underlying principles for monitoring and evaluation and set out its intended approach following the adoption of FP4 ⁽²⁾. This, however, did not constitute a comprehensive evaluation strategy as set out above (see paragraphs 38 and 39). Furthermore, from 1996 to 2006, the Commission did not fundamentally re-examine its approach to evaluation for FP5 and FP6, despite the fact that FPs changed both in scope and orientation (see paragraph 7). In particular, the 'research DGs' have not agreed on a consistent approach to the evaluation of the FPs and its constituent programmes beyond the legislative requirements adopted in 1999 (for FP5), in 2002 (for FP6) and 2006 (for FP7) ⁽³⁾.

42. For the current programming period, the absence of such a strategic approach to evaluation in this field is a particular problem in view of the joint implementation of FP7 by several DGs, the ERC and executive agencies, and the mandatory staff rotation on sensitive posts imposed by the Financial Regulation. It also stands in contrast with current practice in the area of external aid (which is the second largest area of direct management by the Commission). In this area the competent DGs have, since 2002, agreed on common multiannual evaluation strategies linked to programming periods (see paragraph 52).

43. This absence of a clearly formulated evaluation strategy also contrasts strongly with the situation in the USA and Canada, where having a strategy is required by the regulatory framework (see *Annex III*).

44. As illustrated below, there are significant differences in the coverage and intensity of evaluation across the FPs. There were also other weaknesses, such as coordination problems, differences in methodological approaches, timing problems and weaknesses in the communication and follow up of evaluation findings, discussed in later sections of this report. This reflects the absence of a comprehensive evaluation strategy for the FPs.

⁽¹⁾ European Commission, 'Communication from the Commission to the Council on a Community plan of action relating to the evaluation of Community research and development programmes', COM(83) 1 final, 19.1.1983.

⁽²⁾ Communication from the Commission to the Council and the European Parliament, 'Independent external monitoring and evaluation of Community activities in the area of research and technological development', 22.5.1996 (COM(1996) 220 final).

⁽³⁾ Each DG implementing the FPs documents its planned evaluations in annual (and multiannual) evaluation plans. However, these planning documents cannot be said to represent an evaluation strategy as such.

Differences in the coverage and intensity of evaluation across the FP

45. FPs are composed of a number of specific programmes (and sub-programmes) that may differ significantly in terms of their purpose, underlying dynamics and budgetary appropriations (for FP7, see *Annex IV*). Taken together, the specific provisions of the various FP decisions and the Financial Regulation ⁽¹⁾ (which requires the evaluation of programmes and activities which entail significant spending) imply that a legal requirement to evaluate both the FP as a whole and most of its constituent specific programmes has been in force since FP4.

46. The Commission has not established a rule for when specific FP programmes or sub-programmes are to be evaluated or a common threshold above which this must be done. This has led to inconsistent interpretations as to what should be evaluated, and at what level of detail:

- DG Information Society considered the IST programme to be a major spending programme in its own right and organised a full 5YA in 2004 in parallel with that carried out for the FP as a whole,
- DG Research did not systematically organise such an assessment at specific programme level despite the budgets of several programmes under its responsibility being of a comparable size.

47. As a result, for the 2004 5YA, significant parts of FP4 and FP5 programmes were not evaluated at all (corresponding to 30 % of all budgetary appropriations for FP4 and 50 % for FP5). This means that the Commission did not comply with the evaluation requirements specified in Article 27 of the Financial Regulation.

48. These divergent interpretations with respect to the minimum legal requirements regarding evaluation, and the resulting under-evaluation of FPs in the past, illustrate the lack of coordination and the effect of the absence of a strategic approach by the Commission to the evaluation of the FPs.

Recommendations

49. The 'research DGs' should urgently develop and agree on a common evaluation strategy for the FPs. Without being overly prescriptive, this strategy should entail a consistent approach with regard to the evaluation of the FPs and its constituent programmes. Such a strategy should also provide guidance with regard to the minimum level at which detailed evaluation must take place so as to take account of the specificities of each scientific field.

Organising evaluation for the FPs

50. The evaluation of the FPs needs to be supported by appropriate organisational arrangements, which should follow from the Commission's strategic orientation for the evaluation of the FPs. The Court checked the extent to which the organisational structure ensured that the evaluation system was appropriately coordinated, adequately resourced and also perceived by stakeholders as being sufficiently independent.

Commission

FPs are jointly implemented by several DGs whilst responsibility for evaluation is decentralised

51. Within the Commission, responsibility for evaluating expenditure programmes rests with the individual DGs (see paragraph 8) and evaluation units have been set up within each DG. These evaluation units are in charge of planning, coordinating and executing evaluations.

52. The absence of a specific coordinating body or a single entity for the evaluation of the FPs as a whole, such as a joint evaluation office for the 'research DGs', stands in contrast to other policy areas. In the area of external aid, a joint evaluation unit has been established for three DGs (Development, External Relations and EuropeAid) to take charge of the evaluation of the Commission's development programmes in third countries (see paragraph 42).

53. DG Research, via the Commission's internal 'RTD evaluation network', ensures an exchange of knowledge between the evaluation units of the different DGs, but does not have the authority to set minimum evaluation requirements. This mechanism has, however, been insufficient to ensure an appropriate coordination of the Commission's evaluation of the FPs.

54. The Commission revised its general approach to evaluation in 2000 and 2002, and again in 2007 (see paragraph 14). Although no major improvements were observed in the way evaluations of the FPs and their constituent programmes were planned, designed and implemented, in general terms, the 'research DGs' have complied with the formal Commission-wide requirements.

⁽¹⁾ Regulation (EC, Euratom) No 1605/2002, Article 27(4).

The Commission's central services: standard setting, but no enforcement role

55. The Commission's central services (in particular DG Budget and the Secretariat-General) provide methodological advice, define common standards within the Commission as a whole and monitor how the DGs and services apply the standards (see paragraphs 14 and 70). They cannot impose how evaluations are to be done in practice.

56. Experience in the USA and Canada indicates that enforcement by standard-setting bodies is instrumental there in ensuring the effective functioning of the evaluation system. In both countries, federal departments establish and enforce the framework within which a programme's success is ultimately judged (see *Annex III*). Whilst the background is different, their systems are nevertheless based on the same underlying principle that a body separate from the one implementing the programme should assess the design of the evaluation system in some detail and scrutinise its effectiveness. The same can be observed in some Member States (e.g. France ⁽¹⁾, Spain ⁽²⁾, Germany ⁽³⁾), where external bodies supervise evaluations carried out by funding agencies or even carry out these evaluations on their own.

Panels

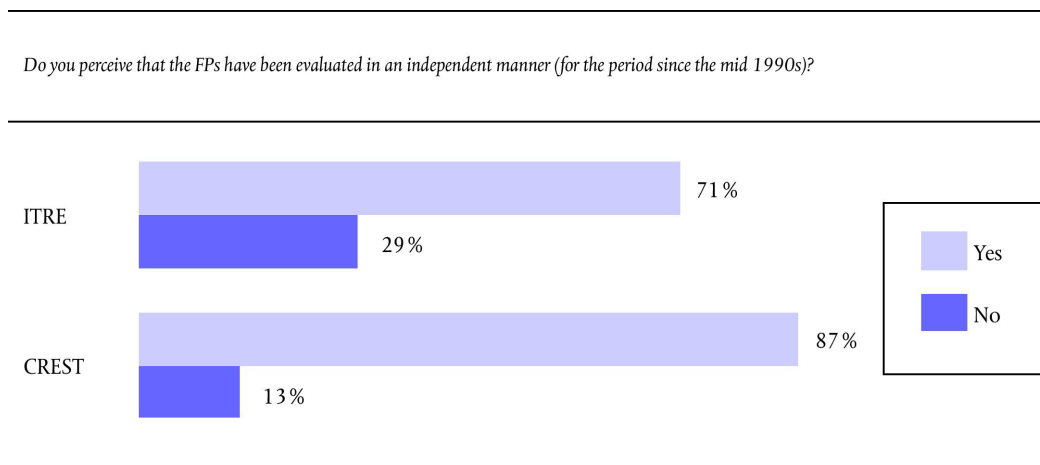
Panels of external experts: required by FP legislation since FP4

57. Since FP4, all major evaluation exercises (and annual monitoring) of the FPs have had to be carried out by (or with the assistance of) panels of external experts (see *Annex II*). Despite difficulties in terms of finding appropriate experts, such panels have the potential to provide transparency, offer a degree of independence, facilitate the involvement of stakeholders and bring to bear knowledge of the FPs themselves, evaluation and the underlying scientific issues.

58. This was confirmed by the Court's surveys, which indicated that respondents saw the Commission's panel-based evaluation system as providing a guarantee of independence (see *Figure 4*).

Figure 4

Perception of institutional stakeholders — independence of panel-based evaluation system



Source: ECA survey 2007.

- (1) 'Loi de programme pour la recherche n° 2006-450' dated 18.4.2006 and 'Décret n° 2006-1334 relatif à l'organisation et au fonctionnement de l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES)' dated 3.11.2006.
- (2) Agencia Nacional de Evaluación y Prospectiva under the Ministerio de Educación y Ciencia: Orden ECI/1252/2005 of 4 May 2005 which creates the 'Comisión Asesora de Evaluación y Prospectiva' (Advisory Commission of Evaluation and Foresight).
- (3) In 2004 the 'Deutsche Forschungsgemeinschaft' (DFG) has created a specialised institute, the 'Institut für Forschungsinformation und Qualitätssicherung' (IFQ), and has vested therein the responsibility for developing and standardising evaluation methodologies, quality assurance and carrying out assessments of the RTD activities funded by DFG.

Weaknesses of the panel approach as implemented by Commission

59. Despite this positive perception, as panel members and stakeholders pointed out to the Court during the audit, the Commission's use of panels for the evaluation of FPs has been characterised by the following weaknesses:

- evaluation panels are usually set up less than a year before their report is due and only convene for a limited number of meetings. This makes it difficult for panel members to influence the definition of the data and analysis required for their work. In addition, because panels are 'ad hoc' and temporary, knowledge about programmes cannot be easily accumulated and relevant issues cannot to be followed up over time. This is a particular weakness when (as for FP7) a number of major evaluation exercises are planned throughout the programming period (see *Annex II*). It also stands in contrast with the Commission's practice for its Expert Advisory Groups (EAGs) for the FPs, which are permanent bodies ⁽¹⁾. Also in Canada ⁽²⁾ evaluation panels are permanent in order to guarantee a consistent and coherent approach (see *Annex III*),
- when defining the terms of reference for panels, and specifying the evaluation questions, the Commission does not consult with external stakeholders, such as industry and major research organisations participating in the FPs. The 2004 monitoring panel has also observed this problem noting that '... potential users should be involved in the process at an early stage' ⁽³⁾, and
- panels operate in a way that is different from 'peer review'. Their members do not have the time to carry out evaluations of their own; they typically provide advice based on data and evidence made available to them from other sources. As a result of the timing constraints, panels have had only limited opportunities to commission additional studies needed to address the questions specified in their terms of reference. In practice, with few exceptions, only information provided by the Commission was available to the 5YA panels.

⁽¹⁾ Under FP7, 15 EAGs have been set up by the Commission, as compared to 12 EAGs under FP6 and 17 EAGs under FP5. These EAGs provide advice to the Commission on policy and programme issues related to the FPs. It should be noted, however, that these EAGs do not completely map into the programme structure of the FPs. In other words, there is not an EAG for each specific programme (or sub-programme).

⁽²⁾ In Canada nearly all federal government departments have established permanent evaluation committees, involving senior departmental officials and usually also external stakeholders, to oversee the evaluation system ('Social Sciences and Humanities Research Council of Canada': Evaluation Standing Committee — Mandate and Membership).

⁽³⁾ European Commission, 'Monitoring Report 2004: Implementation of Activities under the EC and Euratom Framework and Corresponding Specific Programmes', August 2005, p. 5.

Resources*No resources specifically earmarked for the FP evaluation*

60. Funding for evaluation is provided on an annual basis, according to needs specified by the DGs implementing the FPs. It is currently sourced from a combination of Commission-wide resources from DG Budget and resources from the administrative part of the budget at the level of the DGs concerned. In addition, specific measures are funded as indirect actions from the operational budget of the FPs. This means that there is no amount specifically set aside for evaluation.

61. In 2006, the Commission-wide average for staff working in the area of evaluation and activity-based management was estimated at approximately 1 % of total staff ⁽⁴⁾. The corresponding figures for monitoring and evaluation activities of the RTD framework programmes at DG Research and DG Information Society were 1,2 % and 1,3 % respectively. In the case of DG Research, more than two thirds of these staff resources were, however, allocated to the operational programme departments rather than to the evaluation unit, and evaluation was only one of many tasks assigned to them.

Funding for FP evaluation lower than recommended by the Commission

62. Resources devoted to evaluation can also be expressed as a percentage of the overall budgetary appropriations allocated to the programme. Figures obtained during the audit indicate that the practice for research programmes within the EU Member States lies between 0,1 % and 2 %. The Commission's guidelines suggest that 0,5 % of a programme's budget should be devoted to evaluation ⁽⁵⁾. However, for larger programmes, economies of scale mean that a smaller share of their budgets could suffice.

63. A comparison between the Commission's figures for expenditure on evaluation in the 2003-2006 period and the budgetary appropriations for each Directorate-General shows that the overall average runs at around 0,15 %, which is significantly below the level suggested by the Commission. A closer analysis for the year 2006 reveals that DG Information Society spent approximately four times more than DG Research on evaluation compared to the appropriations under their respective responsibilities. In addition, at DG Research, only 0,7 % of one full time official is devoted to evaluation per million euro annual budget. By comparison, DG Information Society devotes double that amount.

⁽⁴⁾ European Commission, 'Planning and optimising Commission human resources to serve EU priorities', SEC(2007) 530, 24 April 2007.

⁽⁵⁾ European Commission, Communication for the Commission from the President and Mrs Schreyer, 'Evaluation standards and good practice', C(2002) 5267, 23.12.2002; European Commission, Staff Working Paper, 'Impact assessment and *ex-ante* evaluation', SEC(2005) 430, 6 April 2005, Annex 1, footnote 209.

64. As illustrated in this report, this additional effort devoted by DG Information Society has resulted in a more comprehensive evaluation coverage (see paragraphs 13 and 46) and, as shown in the sample of evaluation studies reviewed by the Court, the use of methodologies that are better suited to addressing evaluation needs (see paragraph 84).

65. A cost estimate made by the Court based on data provided by the Commission shows that an evaluation system based on the proposals made in this report would, on average, represent less than 0,4 % of the FP7 budgetary appropriations. This would cover the costs of panels, sub-contracted studies, staff of both the joint evaluation office and the current DG-specific evaluation units and other miscellaneous costs ⁽¹⁾.

Recommendations

66. The 'research DGs' should consider setting up a joint evaluation office in addition to the existing evaluation units within each DG. In line with current practice in other policy areas, this office could coordinate the Commission's evaluation activities for the FP as a whole. Such common support functions have been recommended in the past for the DGs implementing the FPs ⁽²⁾ and may become even more important in the future with the creation of the ERC and the delegation of programme implementation to executive agencies.

67. Evaluation panels should continue to be used by the Commission to assist it in its evaluation activities. However, the Commission should consider how to align the panel structure with the overall FP programme structure. This could be done by having panels for specific programmes and, where appropriate, sub-programmes (see *Annex IV*). In addition, these panels should be set up sufficiently early in order to provide effective support to the Commission, and continue to do so throughout the programming period and beyond.

68. The Commission should reconsider the resources to be used for evaluation of the FPs and their constituent programmes.

Evaluation methodologies and techniques

69. Effective evaluation requires the use of appropriate methodologies and techniques. The Court checked the extent to which specific guidelines for the evaluation of the FPs existed, whether information needs were properly analysed on a timely basis and whether relevant data were made available to evaluators. Considering the evaluation studies carried out by (or on behalf of)

⁽¹⁾ In 2007, the European Research Advisory Board (EURAB) also recommended an increase in the resources allocated to the evaluation of the FPs (EURAB, 'Recommendations on Ex-Post Impact Assessment', July 2007).

⁽²⁾ Special Report No 1/2004 on the management of indirect RTD actions under Fifth Framework Programme (FP5) for research and technological development (1998 to 2002) (OJ C 99, 23.4.2004), paragraphs 116 and 122.

the Commission, the Court reviewed whether the techniques used for the individual studies provided a sufficient basis for addressing the evaluation questions set for the 5YAs. The potential contribution that might be made to the evaluation of the FPs by national-level evaluations was also considered.

Deploying FP-specific standards and guidelines

70. DG Budget has issued several Communications setting out its evaluation policy and providing rules for its implementation and has established a set of Commission-wide standards (see paragraph 14 and 55) ⁽³⁾. For other policy areas, the competent DGs have also developed specific evaluation guides ⁽⁴⁾.

71. This is, however, not the case for the FPs. There are a number of methodological guides and toolboxes for evaluation activities in the field of research that have been published (or funded) by the Commission ⁽⁵⁾. However, these neither provide specific guidance on how to carry out evaluations for the FPs nor define how key concepts or terminology are to be used by Commission staff or external experts. This has contributed towards the problems discussed below.

Data needs and data collection

Insufficient analysis of information needs

72. Insufficient attention has been paid by the 'research DGs' to understanding what data is required for evaluation and monitoring purposes. Rather than carrying out a systematic and comprehensive analysis of information needs, the Commission has specified data requirements late in the programming period and without due regard to overall evaluation needs. The FP6 socio-economic questionnaire was the only common attempt among DGs to agree on such data (see paragraph 76). However, this questionnaire was finalised in 2004 (three years after the start of the programming period) and primarily covers aspects such as gender participation and environmental impacts, rather than scientific outputs (such as intellectual property issues, publications, etc.), outcomes or impacts.

⁽³⁾ European Commission, DG Budget, Communication to the Commission from Ms. Grybauskaitė in agreement with the President, 'Responding to strategic needs: reinforcing the use of evaluation', SEC(2007) 213, 21 February 2007; European Commission, Communication for the Commission from the President and Mrs. Schreyer, 'Evaluation standards and good practice', C(2002) 5267, 23.12.2002.

⁽⁴⁾ 14 DGs have such evaluation manuals for the policy areas and programmes under their responsibility (European Commission, DG Budget — Evaluation unit, 'Overview of Evaluation Guides in the Commission', October 2004).

⁽⁵⁾ Fahrenkrog, G. et al, RTD Evaluation Toolbox — Assessing the Socio-Economic Impact of RTD-Policies, IPTS-JRC, 2002; PREST et al; 'Assessing the Socio-economic Impacts of the Framework Programme'; University of Manchester; 2002.

73. The assessment of programmes relies heavily on the availability of three kinds of data:

- data collected according to effectiveness indicators, the analysis of which provides intelligence about progress towards programme objectives. The lack of an intervention logic referred to above means that it has been difficult to define such indicators (see paragraph 30). In practice, the data that were collected were of little relevance for measuring the achievement of programme objectives,
- data about programme management (according to efficiency indicators), and
- data about participants (such as the participation rates of major research organisations).

74. In particular, the latter two should be available at the Commission. However, due to problems with respect to the Commission's internal IT systems, which are the source of most of the data on programme management and participants ⁽¹⁾, significant problems have been experienced with this in practice ⁽²⁾.

75. This is illustrated by the considerable delays in making Article 173 Reports and Annual Monitoring Reports available (up to 16 months late for Article 173 Reports ⁽³⁾ and up to seven months late for Annual Monitoring Reports in the period between 2002 and 2005). As this information was consistently late, it was of limited use.

⁽¹⁾ Whilst the Commission generally does have such information regarding participants, it often cannot easily be used for detailed analysis because data concerning a given participant is neither kept in a single place nor necessarily recorded in compatible formats. This makes it difficult to track the participant's involvement over time, for example to assess reasons for non-participation, and relations between participants, for example to study networking effects (see paragraph 84).

⁽²⁾ Annual Report concerning the financial year 2006, table 2.3; Annual Report concerning the financial year 2005, paragraph 7.13; Annual Report concerning the financial year 2004, paragraphs 6.8 and 6.10; Annual Report concerning the financial year 2003, paragraphs 6.9 and 6.70; Special Report No 1/2004, paragraphs 105-108; Opinion No 1/2006 on the proposal for a regulation of the European Parliament and of the Council laying down the rules for the participation of undertakings, research centres and universities in actions under the Seventh Framework Programme and for the dissemination of research results (2007 to 2013) (OJ C 203, 25.8.2006), paragraph 57.

⁽³⁾ Treaty establishing the European Community, Article 173: 'At the beginning of each year the Commission shall send a report to the European Parliament and the Council. The report shall include information on research and technological development activities and the dissemination of results during the previous year, and the work programme for the current year.'

Ineffective approach to sourcing evaluation data from participants

76. Since FP5, the Commission has systematically required participants, by contract, to provide data needed to measure the achievement of programme objectives in terms of outputs, outcomes and impacts, as follows:

- under FP5, participants were required to submit a 'Technological Implementation Plan' together with a final report after the project had been completed. These documents were often of poor quality or were not submitted at all,
- under FP6, they are required to produce a 'Plan for Disseminating and Using Knowledge' and must respond to a standardised questionnaire concerning socio-economic issues when submitting their annual technical reports.

77. These contractual requirements to provide data have placed unnecessary administrative burdens on participants. For example, with respect to the FP6 annual questionnaire, data are collected much more frequently than needed and from the entire population of projects when a representative sample would suffice. This is likely to lead to 'evaluation fatigue'. Moreover, this information has been collected at the wrong time; many of the underlying issues for which data are collected are medium or longer-term in nature so that data collection within the contractual duration of the project is by definition premature ⁽⁴⁾. This is why only limited practical use has been made of the information provided by these tools for evaluations under FP5 and FP6.

78. However, none of the systems in place provides for information to be collected at the end of (or shortly after) a project's lifespan in a standardised manner. Stakeholders interviewed suggested that such information might include the achievement of project objectives, intended (and unplanned) results and the perception of Community added value by participants. Such information could also form the basis for external reviews of completed projects ⁽⁵⁾.

79. Together with a more extensive use of scientometric approaches (such as citation analysis) to assess project results (see paragraph 83), this could provide a basis for the Commission to move to a more results-based financing system for RTD activities co-funded by the Community ⁽⁶⁾.

⁽⁴⁾ There has been no systematic attempt to collect data from participants after the completion of the project, except for specific surveys in the framework of evaluations. The use of surveys has been hampered by the fact that participants, unlike in a number of national RTD programmes (e.g. Germany), are not legally obliged to provide such analytical data.

⁽⁵⁾ Special Report No 1/2004, paragraph 94; Opinion No 1/2006, paragraph 56.

⁽⁶⁾ Annual Report concerning the financial year 2006, paragraph 7.33; Annual Report concerning the financial year 2005, paragraphs 7.8 and 7.29; Opinion No 1/2006, paragraphs 58 to 61.

Other data sources inadequately used

80. Insufficient use is made of other existing sources of data within the Commission and the Member States (e.g. databases of national RTD programmes and funding agencies and the 'Community Research & Development Information Service' (CORDIS) operated by the Commission's Publication Office ⁽¹⁾).

Evaluation studies carried out by the Commission*Reports reviewed by the Court*

81. The Court reviewed 86 evaluation studies carried out by (or on behalf of) the Commission since 2000. These included all 36 studies forming the basis for the 5YA published in 2004 and the further 50 studies completed since that date. Of these, only 23 studies can be considered to be genuine evaluations of the FP and its programmes (see Figure 5).

Figure 5

Overview of evaluation studies carried out by (or on behalf of the Commission) in the period 2000 to 2006

Type of studies	Number	Description
Evaluations of FP activities	23	Evaluation studies — based on recognised methodologies ⁽¹⁾ , — make specific reference to programme objectives and to the extent to which programme objectives have been met.
National impact assessments undertaken at Member State level	12	Studies carried out by Member States which address mainly national participation rates, rather than concluding about programme objectives.
Methodological studies	5	Studies organised by the Commission to assist in the development of evaluation methodologies.
Studies of the European RTD landscape	2	Studies addressing contextual analyses on the research scene in Europe, but not the FPs. In particular no explicit reference is made to the effect that the FPs have on national research and innovation systems.
Non-analytical descriptive and statistical documents	33	Documents which typically address FP activities in a very indirect fashion, often describing activities rather than analysing them.
Other documents	11	A range of documents that consider issues of general policy interest, without examining the FPs in that context.

⁽¹⁾ Hanne Foss Hansen, 'Choosing evaluation models: A discussion on evaluation design', *Evaluation*, Vol 11 No 4, 2005, pp. 447-462.

Evaluations focusing on monitoring issues, rather than outcomes and impacts

82. Two main observations result from the Court's meta-analysis of these specific evaluation studies:

- whilst many of the studies attempt to conclude against programme objectives, these conclusions are presented in a vague manner, leaving the reader unconvinced that the

evaluation has actually addressed this issue. This is often because the lack of any clear programme objectives critically undermines the process (see paragraph 28),

- rather than evaluating the outcomes and impacts of the FPs or their specific programmes (and sub-programmes), many of the documents are descriptive, and, at best, concentrate on immediate project outputs and project management matters. These reports are unable to provide convincing evidence with regard to outcomes or impacts.

⁽¹⁾ CORDIS is the only database that provides information on proposals, projects and results that is common to the whole FP, covers several programming periods, and is accessible from outside the Commission. However, data in CORDIS for individual projects are often not updated after the proposal has been selected. Up to and including FP6, public and private sector participants in the FPs were not obliged to make public their results (European Research Advisory Board (EURAB), 'Scientific publication: policy on open access', December 2006).

Insufficient focus on the development of innovative methodologies

83. The use of traditional methodologies and techniques (in particular surveys, document review and interviews) dominates

for evaluations carried out in the period 2000 to 2006 (see Figure 6). Furthermore, a number of other sophisticated, but nevertheless well-established, evaluation methodologies and techniques were not (or only very rarely) employed, in particular:

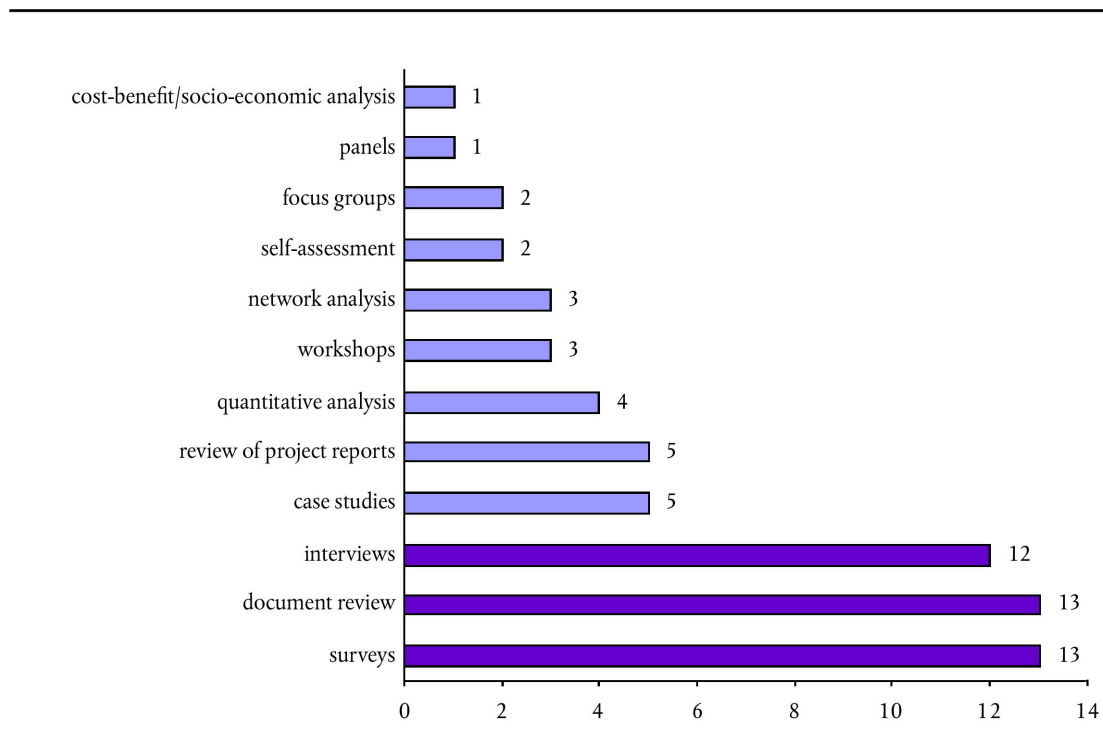
- scientometric approaches (such as citation analysis) ⁽¹⁾,
- quantitative models (such as cost-efficiency, cost-effectiveness and cost-benefit) ⁽²⁾, or
- actor (or behavioural) models (which focus on the perspectives of those involved in an intervention, such as the characteristics of research networks and changes thereto over time).

84. Evaluation studies making an attempt to address some of the methodological difficulties of assessing public intervention in research (see *Annex I*) were used only in a limited number of cases:

- recently two studies were completed on the subject of intervention logic ⁽³⁾,
- only three studies have been identified which attempt to assess the effect that Community funding has had on the pan-European networking behaviour of research organisations ⁽⁴⁾. This is despite the fact that the assumption of positive networking effects is at the heart of the ERA concept,

Figure 6

Use of methodologies and techniques in FP evaluations (2000-2006)



Note: Use of several techniques per evaluation possible.
Source: ECA.

⁽¹⁾ Up to and including FP6, public and private sector participants in the FPs were not obliged to make public their scientific results. Guaranteeing free public access to Community funded research by making publication mandatory has been recommended in December 2006 by the EURAB. It was also suggested that the Commission is to set up a web-based repository for published project results (see EURAB, 'Scientific publication: policy on open access', December 2006).

⁽²⁾ It is noted that such models were used to make forecasts of the micro and macro-economic effects of the planned FP7 (see European Commission, 'Monitoring Report 2004: implementation of activities under the EC and Euratom framework and corresponding specific programmes', August 2005, pp. 23-25). However, there is no evidence that this assessment of the potential economic impacts of the FPs is linked to any plan for *ex-post* evaluation.

⁽³⁾ DG Information Society and Media, 'IST-RTD Monitoring and Evaluation: Results and Impact Indicators (2003-2004)', 2005; 'Networks of innovation in information society development and deployment in Europe', 2006.

⁽⁴⁾ DG Information Society and Media, 'ERAnets — Evaluation of Networks of Collaboration among participants in IST research and their evolution to collaborations in the European Research Area' (Final Report), 2005; 'Evaluation of progress towards a European Research Area for information society technologies', 2005; 'Networks of innovation in information society development and deployment in Europe', 2006.

- in only one case has an attempt been made to draw conclusions about the impact achieved at programme (or sub-programme) level by combining the aggregation of project level evaluation analyses with other evaluation techniques ⁽¹⁾, and
- no evidence has been found for the use of other methodologies, such as 'counterfactual' (or 'control group') approaches. This is an approach that would be particularly useful for illustrating the added value of programmes at Community level compared to not intervening at all (or intervening at national or regional level).

85. Stakeholders interviewed by the Court were of the opinion that the usefulness and relevance of the Commission's evaluations could be improved. This demonstrates the need for further methodological development so that the evaluation system can

evolve to take account of the changing character of the FPs (see paragraph 7) ⁽²⁾.

Challenging issues raised in 5YAs could not be addressed on the basis of the Commission's specific evaluation studies

86. This is also illustrated by the significant problems faced by the 2004 5YA panel ⁽³⁾ in answering the evaluation questions suggested in the terms of reference. Many of these challenging questions could not be examined at all, due to the absence of appropriate studies carried out by (or on behalf of) the Commission beforehand and the impossibility of the panel members to commission such studies in good time (see Figure 7).

Figure 7

Addressing evaluation questions in the 2004 5YA

	Evaluation question	Addressed in report?
IMPLEMENTATION	Were the activities carried out efficiently?	partially
	Were the activities cost effective?	partially
	Did the activities constitute the best way of achieving the objectives set?	no
	Were the overall legal framework (including rules for participation and contracts), policy instruments and the modalities for implementation clear, appropriate and effective?	no
	Were the level of funding and other available resources adequate?	no
	Were the targeted industrial and research communities, including SMEs, able to respond appropriately?	yes
ACHIEVEMENTS	Did the activities attain their respective objectives and to what extent were there unexpected results?	no
	What are the major results in particular in terms of scientific, technological, socio-economic and environmental outputs, in terms of international co-operation, knowledge transfer and innovation, pre-normative activities, accessibility, dissemination and uptake of research, human resources development, mobility and training, and in terms of supporting and enhancing co-ordination of research activities?	no
	Were the results and their effects and impacts globally satisfactory from the point of view of direct or indirect beneficiaries and stakeholders?	no
	Were the relevant industrial and research communities, including SMEs, addressed satisfactorily?	yes
	How and how far have the activities contributed to improved EU research competitiveness at international level? Did EU research attain leadership in specific areas?	partially
	How and how far have activities contributed to EU policies in general and to the EU's strategy for sustainable development?	no
	Is there evidence of a structural change, including in particular networking, integration and coordination of research, at a national or at an international level as a result of Community research activities?	no
	Did the programmes provide value for money? Did the activities have lasting impacts?	no
EVOLUTION AND FUTURE PERSPECTIVES	How did Framework programmes evolve in terms of rationale, objectives, thematic priorities, balance between bottom-up and top-down priorities and between fundamental and applied research, instruments, European added value and other impacts, in particular in the light of the emergence of the European Research Area (ERA) concept, the Lisbon objectives and major international economic and research benchmarks?	yes
	How did the Framework Programme achieve European added value?	yes
	Can European added value be achieved through other international or national mechanisms?	no
	Are the programmes' policy objectives, priorities, instruments and lifecycle appropriate for the future?	yes

⁽¹⁾ DG Information Society and Media, 'Framework Contract for Impact Assessments for Research in the Field of Information Society Technologies'.

⁽²⁾ EURAB, 'Recommendations on Ex-Post Impact Assessment', July 2007.

⁽³⁾ European Commission, 'Five-year Assessment of the European Union Research Framework Programmes, 1999-2003'. Report of the Independent Expert Panel chaired by Erkki Ormala (15 December 2004), Annex 1.

87. In addition, it is noted that only for those aspects where clear and quantifiable objectives were defined in the programmes (such as the share of small and medium-sized enterprises participating), were the panel members able to answer the questions specified in the terms of reference.

Evaluations carried out at national level for the FPs

88. Several Member States and associated countries carry out their own national evaluations of the FPs, which address relevant and interesting questions from their own perspective ⁽¹⁾. In the last eight years, 14 such studies have been produced for nine Member States and one associated country.

89. Whilst the Commission encouraged standardised approaches where possible, these national impact studies are very heterogeneous ⁽²⁾. As a result, since they do not follow a harmonised methodology, possibilities to aggregate, combine or compare results are limited. Moreover, given their national character, these studies can only partly capture the European cooperative dimension.

90. In the 2007 ERA 'Green paper' the Commission argued that it was important that accurate analyses should be available at ERA level on the impact and effectiveness of research activities and policies, including those of the Member States and the EU. It also recognised that current arrangements were not adequate to address this challenge because each level of intervention performed evaluations separately from the others ⁽³⁾.

Recommendations

91. The Commission should draw up a manual for the evaluation of activities funded by the FPs, providing definitions of key concepts and terminology and identifying a range of appropriate methodologies for certain evaluation activities. This manual should be based on existing guides and toolboxes and should be updated on a regular basis.

92. The Commission should define its information needs, taking into account the need to limit the administrative burden on participants.

93. The Commission should provide monitoring information on the implementation of the FPs in a timely manner. As required by the FP7 'Rules for Participation', problems relating to the Commission's IT systems for managing data on programme management and participants should be addressed urgently.

⁽¹⁾ European Commission, 'The impact of EC R&D policy on the European science and technology community; National impact studies synthesis final report' (1995).

⁽²⁾ European Commission, 'Design of a harmonised impact study methodology for the R&D EU Framework Programmes', 1992.

⁽³⁾ European Commission, Staff working document, 'The European Research Area: new perspectives', SEC(2007) 412/2, 4 April 2007, section 3.3.3 'ERA and the implications for evidence-based monitoring and evaluation', p. 93.

94. Participants' reporting obligations towards the Commission should be reduced significantly and surveys carried out on a sample basis only. The contractual requirement to submit indicator-based impact data for each project should be replaced by a short self-assessment upon project completion.

95. The Commission should make more extensive use for its evaluations of existing complementary sources of relevant data (both within the Commission and in Member States).

96. The Commission should employ the full range of evaluation techniques available. In addition, it should support the development of innovative approaches to address the methodological difficulties inherent to evaluating public intervention in research (see *Annex I*) and use these approaches in its own evaluations.

97. The Commission should consult with Member States on how to maximise the benefits of national-level evaluations of the FPs.

Disseminating and using evaluation findings

98. Evaluation should provide relevant information and analysis that can be effectively used for programme management and policy making. In particular, evaluation should provide the basis for the informed reflection on the strengths and weaknesses of a public intervention that is necessary for learning to take place. The Court checked the extent to which evaluations were properly timed, findings were communicated and disseminated to stakeholders outside the Commission and recommendations were taken into account by the Commission to revise ongoing programmes (learning programmes).

Timing of evaluations did not take into account the type of information that could realistically have been expected

99. The Court recognises that certain types of analysis require a long-term perspective (e.g. the evaluation of outcomes and socio-economic impacts) and that some aspects are related to a specific programming period (e.g. programme objectives, even within a given scientific field), whereas others are not (see *Figure 8*). However, the Commission has not adopted an approach in its evaluation activities that takes these time horizons into consideration.

100. The legislation has required the Commission's main evaluation activities to focus exclusively on the five years preceding the 5YA thereby covering parts of two successive programming periods. As a result, medium and long-term outcomes and impacts have barely been addressed. As a consequence, the 5YAs have generally focused mostly on monitoring aspects, short-term programme design issues and general observations about the orientation of the FPs. The Court welcomes the fact that FP7 legislation no longer requires 5YAs, or any other pre-defined period of coverage.

Figure 8

Monitoring and evaluation issues according to different time horizons

Monitoring of programme implementation (on-going)	— <u>Programme management issues</u> (e.g. time-to-contract, time-to-payment, subscription and success rates, budgetary execution, funding rates, etc.)
Evaluation Short term (say, at the earliest after 1-2 years)	— <u>Programme management issues</u> (e.g. efficiency of administrative procedures) — <u>Programme design issues</u> (e.g. accessibility and flexibility of instruments, barriers to participation, implications of non-success in calls for proposals, etc.) — <u>Analysis of participation profile</u> (mainly quantitative at this stage)
Evaluation Medium term (say, after 7 years)	— <u>Programme design issues</u> (e.g. effectiveness of instruments) — <u>Analysis of participation</u> (combining qualitative and quantitative aspects, e.g. research networks analysis, impacts on the behaviour of researchers, etc.) — <u>Analysis of project outputs</u> ⁽¹⁾ (e.g. publications, conference papers, patents, licences, prototypes, standards, trained scientists, etc.)
Evaluation Long term (say, after > 10 years)	— <u>Assessment of outcomes and impacts</u>

⁽¹⁾ It should be noted that grant agreements for projects are signed throughout the programming period, the first agreements generally coming into force at the end of the first year of a programme. The duration of these agreements is mostly three to five years, as a result final outputs can be expected at the earliest after four to six years.

Source: European Court of Auditors.

101. The shortcomings of the 5YAs in assessing the outcomes and impacts of the FPs has also been noted by the Commission's experts:

- the 2004 5YA panel noted that '... evaluation should cover and give sufficient attention to both long-term and short-term issues' and '... resources could be released [...] to allow more ambitious long-term evaluation studies that could explore the structural and wider socio-economic issues in a more systematic way' ⁽¹⁾,
- a study commissioned by DG Research found that '... current EU RTD evaluation practices (comprising continuous monitoring, five year assessments and mid term evaluation) are characterised by strong focus on monitoring compared to impact assessment' and '... in order to shift demand and expectation [...] towards socio-economic impact there would need to be a new evaluation and monitoring framework that obliged a greater effort at impact level rather than monitoring of outputs' ⁽²⁾.

⁽¹⁾ European Commission, 'Five-year assessment of the European Union Research Framework Programmes, 1999-2003', Report of the Independent Expert Panel chaired by Erkki Ormala (15 December 2004); Section 6 'Evaluating the Framework Programme', p. 19.

⁽²⁾ European Commission, 'Evaluation Toolbox — Assessing the Socio-Economic Impact of RTD Policies', the so-called 'EPUB Toolbox' edited by Joint Research Centre (JRC) — Institute for Prospective Technological Studies (IPTS) and Joanneum Research (August 2002), Executive Summary, pp. 9 and 12.

102. The issue of timeliness also affects specific evaluations required by the legislation. For example, the 2004 Marimón report, which analysed the new instruments introduced in FP6 (2003-2006), was required too early, and as a result, there was little or no practical experience and empirical evidence on which to base its conclusions ⁽³⁾.

Communication and dissemination

103. On its own, the Commission can develop and adapt operational work programmes according to short-term cycles, subject to approval by the relevant programme committee. This could represent an opportunity for adjustments to be made to programme management, and also for reallocations of budgets within sub-programmes. No evidence was found that the Commission makes use of this possibility (see paragraph 107).

104. Beyond this, given the way the FPs are designed and adopted, it is all but impossible for the Commission to make significant changes, such as modifications of programme objectives or the reallocation of budgets between sub-programmes. Only the legislator can enact such wide-ranging modifications during a revision of the FP legislation (such as in 2010 for FP7) or the adoption of the subsequent FP. For this reason, the legislation has typically stipulated that the main evaluation exercises be timed so that their results could be most usefully taken into account for policy making.

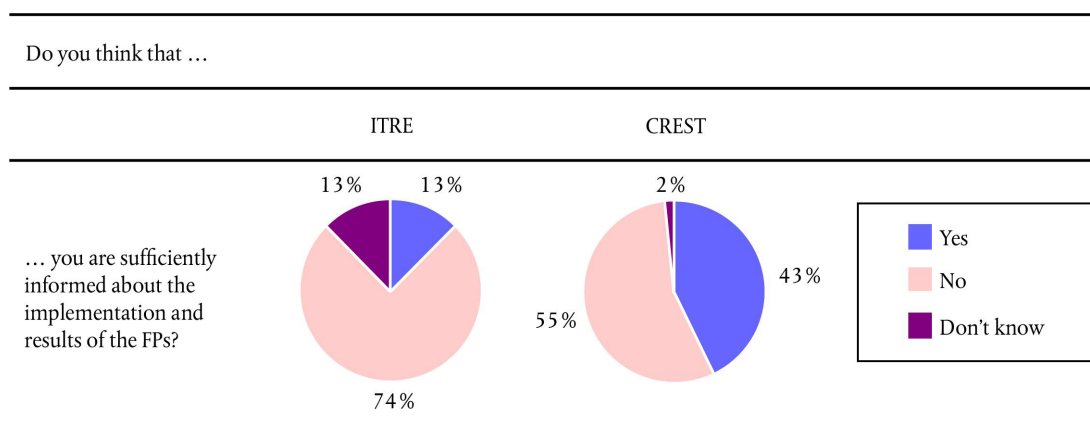
⁽³⁾ European Commission, 'Evaluation of the Effectiveness of the New Instruments of Framework Programme VI', 21 June 2004, Report of a High-level Expert Panel chaired by Professor Ramon Marimón, p. 20.

105. Therefore, some of the main users of the Commission's evaluations are the 'research DGs' themselves ⁽¹⁾. However, the FP legislation also explicitly requires evaluations to be communicated to the European Parliament and the Council.

106. In the Court's survey of the European Parliament's ITRE committee and the Council's CREST committee, a majority of respondents, however, indicated that they did not feel sufficiently informed about the implementation and the results of the FPs (see Figure 9) ⁽²⁾.

Figure 9

Perception of institutional stakeholders — satisfaction with information received



Source: ECA survey 2007.

Follow-up of evaluation findings

107. In its Annual Evaluation Reviews, the Commission provides a summary of findings and actions taken stemming from its evaluations ⁽³⁾. With regard to the follow-up, the Commission has also provided formal replies to the findings and recommendations of the 5YAs and the specific evaluations required by the legislation and, to the extent possible, has taken appropriate measures to address the weaknesses identified therein. This was, however, not the case for many of the other evaluation reports reviewed and no evidence was found that their findings and recommendations were taken into account for amendments to work programmes. Similarly, the DGs' ABB budgetary statements and their Annual Activity Reports do not indicate the extent to which evaluation findings were acted upon.

Recommendations

108. The Commission should establish the type and scope of evaluation that can reasonably be expected for the dates specified in the legislation (2008, 2010 and 2015), explain how evaluations can be used to adapt programmes (learning programmes) and specify what contribution evaluations can make to policy decisions. In particular, the Commission should clarify its intended approach with regard to the evaluation of the longer-term results of past and present FPs, given that such results may take seven to 10 years to come to fruition (see Figure 8).

109. The Commission should review its practice on communicating and disseminating its evaluation findings to ensure that stakeholders receive the necessary information. In particular, the 'research DGs' should provide a formal response to all evaluations and make these studies public in a timely manner, together with their replies. The 'research DGs' should also report in their ABB budgetary statements and their AARs on the follow-up given and make use of recommendations when updating work programmes and designing future programmes.

⁽¹⁾ European Policy Evaluation Consortium (EPEC), Technopolis France, 'Study on the use of evaluation results in the Commission: Final report', Dossier No 1 Synthesis report and annexes, May 2005.

⁽²⁾ The higher level of satisfaction reported by respondents from CREST, who are often also members of Programme Committees, may be due to their formal role in monitoring programme implementation.

⁽³⁾ European Commission, Staff working paper, Annual Evaluation Review 2004 — Overview of the Commission's evaluation findings and activities, SEC(2005) 587, 28 April 2005.

OVERALL CONCLUSION

110. The Court recognises that evaluating RTD programmes, and in particular assessing their long-term results, is inherently difficult and best practice is hard to define.

111. In the Court's view, evaluation is essential for programme management, for future planning and for policy making. Evaluation also provides a basis for accountability. It provides valuable information about what intended (or unintended) results have been achieved, the extent to which programme objectives have been reached, the relevance and utility of programmes and the efficiency and effectiveness of programme implementation. As it is the European Parliament and the Council that decide upon general orientation, budgetary matters and programme design for the FPs, a transparent, credible and robust evaluation system is needed to provide them with the appropriate information.

112. The Commission established an overall evaluation framework in 2000 and 2002 and subsequently revised it in 2007. Although during the period audited by the Court no major improvements were observed in the way evaluations of the FPs and their constituent programmes were planned, designed and implemented, in general terms, the 'research DGs' have complied with the formal Commission-wide requirements.

113. For the period covered by this report, as required in the FP legislation, the 'research DGs' had an evaluation system in place and the Commission can point to a sizeable body of evaluation studies. However, the fact remains that little or nothing is known about the achievement of programme objectives and the results of the FPs. This is because evaluations have generally focussed on short-term issues of programme implementation. As a result, the Commission's evaluations were of limited usefulness to policy-makers, stakeholders and even to the Commission itself.

114. Given the importance of evaluation for programme management and policy-making, the Commission's approach to evaluating the FPs and their constituent programmes needs to be re-examined in view of new political challenges, increased funding, a broadening of the orientation of Community research policy and the recent Commission reform. This was also recommended by EURAB in its 2007 report.

115. In the Court's view, the use of an explicit intervention logic in future FP programme design would lead to more focussed and better structured programmes. It would also contribute to the Commission's 'better regulation' policy, which aims for a better designed, simpler, more effective and better-understood regulatory environment ⁽¹⁾.

116. Evaluation of the FPs and their constituent programmes would be more effective if programme objectives and indicators were clearer from the outset, a robust, but not overly prescriptive, evaluation strategy were agreed upon by the 'research DGs' implementing the FPs and more advanced evaluation methodologies, including quantitative approaches, were used.

117. The Court also noted a number of respects in which there is scope for the Commission to reconsider its existing organisational arrangements for the evaluation of the FPs. There may be advantages in:

- involving evaluation panels earlier in the evaluation process, thereby providing assistance to the Commission throughout the programming period and beyond, and increasing specialisation by having panels for specific programmes and, where appropriate, sub-programmes,
- supplementing the existing DG-specific evaluation units with a joint evaluation office in charge of coordinating the FP evaluations.

118. The recommendations contained in this report are based on either existing practices within the Commission or international good practice. With regard to resources, it is clear that the introduction of the suggested approach to programme evaluation would entail higher costs. However, in the Court's view, it is most likely that this will produce a positive pay-off of at least the same order.

This report was adopted by the Court of Auditors in Luxembourg at its meeting of 22 November 2007.

For the Court of Auditors
Hubert WEBER
President

⁽¹⁾ European Commission, DG Budget, Communication to the Commission from Ms Grybauskaitė in agreement with the President, 'Responding to strategic needs: reinforcing the use of evaluation', SEC(2007) 213, 21 February 2007, p. 4.

ANNEX I

MEASURING THE IMPACT AND RESULTS OF PUBLIC INTERVENTION IN RTD

Some of the main methodological difficulties in measuring the impact and results of public intervention in RTD are:

- **‘attribution problems’:** attribution problems are manifest in evaluations and performance measurement. They are closely linked to the ‘additionality’ question, which asks whether changes would not have happened anyhow. The difficulty of attributing results to a specific public intervention is also influenced by a large number of external factors (such as the effects of other policies or programmes, changes in the legal framework, societal changes, etc.). Furthermore, in research, not achieving the expected results can also be a success. A further complication is the apparent difficulty of clearly attributing research outputs, not to mention outcomes and impacts, to specific aspects of the RTD process, or even to individual entities involved in a research activity,

- **‘measurement problems’:**
 - **‘data availability’:** generally speaking, most data available, in particular the data collected through monitoring systems, is on the inputs and outputs of RTD activities, rather than on outcomes or impacts,

 - **‘understanding the dynamics’:** given the apparent difficulty to model the causal relations between inputs and outputs, the inner workings of the research process are unclear; and in the absence of an explicit intervention logic, the process itself is mostly considered as a ‘black box’,

 - **‘comparability of results’:** the extent to which a comparison of results between different scientific fields can be made is limited, so that these generally should be interpreted in the context of specific programmes and sub-programmes. In addition, the evaluation of results is more difficult for basic (or fundamental) research, than for applied research or technological development. The main reason for this is the fact that basic research generally consists of unique, non-replicable procedures. Even in the field of applied research, results vary considerably and cannot therefore be compared easily,

 - **‘aggregation’:** concluding from evaluation results in terms of outputs, outcomes and impacts attained at the lower level for the next higher level poses problems (for example when aggregating project results, first for clusters of projects, then for domains within a programme and finally at the level of the programme as such),

 - **‘adequacy of indicators’:** another question is whether evaluators are measuring the right thing. Metrics must be easy to measure, difficult to manipulate and drive the right behaviour. However, a ‘measurable’ indicator for output does not necessarily permit conclusions on outcomes or impacts to be drawn,

- **‘timing problems’:** a further obstacle is the considerable time lag from the research activity to the generation of outputs and outcomes, so that impact can be assessed. Data needs to be collected over the long term so that meaningful and robust conclusions can be drawn. Such time scales for evaluation are often not consistent with the expectations and needs of policy makers, and a trade-off must be made between early delivery for policy making and allowing the adequate time for the effects to emerge,

- **‘project fallacy’:** evaluation of outcomes and impacts of individual projects within a programme needs to take account of the dilemma that the technological and scientific results have usually been obtained using previous work. Similarly, before the publication and exploitation of the scientific results takes place, the project results may be used for other work by the participant. Therefore, a project in itself is not necessarily the appropriate assessment framework, in particular for longer-term outcomes and impacts,

-
- **‘credibility, integrity and independence’**: the quality of an evaluation depends largely on the level of engagement the evaluators have with the subject matter, which involves a trade-off between understanding and expertise on the one hand, and independence and objectivity on the other hand. It also depends on the assumptions and interests of policy makers and the rationales and purposes of the evaluation.

ANNEX II

PROVISIONS REGARDING MONITORING AND EVALUATION IN LEGISLATIVE DECISIONS FOR FP4 TO FP7

FP4 (1995 to 1998) ⁽¹⁾	FP5 (1999 to 2002) ⁽²⁾	FP6 (2003 to 2006) ⁽³⁾	FP7 (2007 to 2013) ⁽⁴⁾
Framework programme decision — Council and EP			
Decision No 1110/94/EC of the European Parliament and the Council of 26 April 1994 concerning the Fourth Framework Programme (1994 to 1998) in the field of research and technological development and demonstration: Article 4	Decision No 182/1999/EC of the European Parliament and of the Council of 22 December 1998 concerning the Fifth Framework Programme of the European Community for research, technological development and demonstration activities (1998 to 2002): Articles 5 and 6	Decision No 1513/2002/EC of the European Parliament and of the Council of 27 June 2002 concerning the Sixth Framework Programme of the European Community for research, technological development and demonstration activities, contributing to the creation of the European Research Area and to innovation (2002 to 2006): recital (20), Articles 4 and 6, Annex III	Decision No 1982/2006/EC of the European Parliament and of the Council of 18 December 2006 concerning the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013): Article 7
Monitoring			
— to continually and systematically monitor, with appropriate assistance from independent external experts, the progress in relation to its initial objectives (Article 4(1))	— to continually and systematically monitor each year, with the help of independent qualified experts, the implementation of the FP and its specific programmes in the light of criteria specified in Annex I and the scientific and technological objectives specified in Annex II (Article 5(1))	— to continually and systematically monitor (each year) with the help of independent qualified experts, the implementation of the FP and its specific programmes (Article 4 and 6(1))	— to continually and systematically monitor the implementation of the FP and its specific programmes (Article 7(1))
	— to assess whether the objectives, priorities and financial resources are still appropriate to the changing situation, and if not to submit proposals on how to adapt or supplement the FP and/or the specific programmes (Article 5(1))	— to review regularly all research activities including, in particular, the monitoring of the level of research excellence (Annex III, 1. Instruments)	
			— to regularly report and disseminate the results of this monitoring (Article 7(1))
Article 173 — Annual report			
		— to report in detail on progress with implementing the FP, and in particular progress towards achieving its objectives and meeting its priorities as set out under each heading of Annex I; including information on financial aspects and the use of the instruments in the context of the annual report to be submitted by the Commission pursuant to Article 173 of the Treaty (Article 4)	

FP4 (1995 to 1998) ⁽¹⁾	FP5 (1999 to 2002) ⁽²⁾	FP6 (2003 to 2006) ⁽³⁾	FP7 (2007 to 2013) ⁽⁴⁾
Interim assessment			
	<ul style="list-style-type: none"> — to review progress with the programme <u>half-way through the term of the FP</u> (Article 6) — to submit to the EP and to the Council, on the basis of the assessments of the various specific programmes, a <u>communication</u> accompanied, if appropriate, by a proposal for the adaptation of this Decision (Article 6) 		<ul style="list-style-type: none"> — to carry out by 2010, with the assistance of external experts, <u>an evidence-based interim evaluation</u> of this FP and its specific programmes building upon the <i>ex-post</i> evaluation of the previous FP, covering in particular the quality of the research activities under way, the quality of implementation and management, and progress towards the objectives set (Article 7(2)) — to <u>communicate</u> the conclusions thereof, accompanied by its observations and, where appropriate, proposals for the adaptation of the FP, to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (Article 7(2))
	<ul style="list-style-type: none"> — to <u>regularly inform</u> the European Parliament and the Council of the <u>overall progress</u> of the implementation of the FP and the specific programmes (Article 5(4)) 	<ul style="list-style-type: none"> — to submit <u>regular progress reports</u> to the European Parliament and the Council on the implementation of the FP (recital (20)) 	<ul style="list-style-type: none"> — to present a <u>progress report</u> (proceeding the interim evaluation) as <u>soon as enough data becomes available</u>, giving initial findings on the effectiveness of the new actions initiated under FP7 and of the efforts made on simplification (Article 7(2))
		<ul style="list-style-type: none"> — to <u>undertake in 2004</u> an evaluation by independent experts of the effectiveness of each of the instruments in the execution of this programme (Annex III, 1. Instruments) 	
Five-year assessment (5YA)			
<ul style="list-style-type: none"> — to have an external assessment conducted by independent qualified experts into the management of and progress with Community activities carried out <u>during the five years preceding this assessment</u> (Article 4(2)) 	<ul style="list-style-type: none"> — to have an external assessment conducted by independent highly qualified experts of the implementation and achievements of Community activities carried out <u>during the five years preceding this assessment</u> in the light of the criteria set out in Annex I, the scientific and technological objectives set out in Annex II (Article 5(2)) 	<ul style="list-style-type: none"> — to have an external assessment conducted by independent highly qualified experts of the implementation and achievements of Community activities carried out <u>during the five years preceding this assessment</u> (Article 6(2)) (respectively of the implementation of the activities undertaken, bearing in mind the contribution of FP6 to the creation of the European Research Area (recital (20)) 	

FP4 (1995 to 1998) ⁽¹⁾	FP5 (1999 to 2002) ⁽²⁾	FP6 (2003 to 2006) ⁽³⁾	FP7 (2007 to 2013) ⁽⁴⁾
<ul style="list-style-type: none"> to <u>communicate</u> the assessment, accompanied by the Commission's comments, to the European Parliament, the Council and the Economic and Social Committee <u>prior to presenting a proposal for the subsequent FP</u> (Article 4(2)) 	<ul style="list-style-type: none"> to <u>communicate</u> the assessment, accompanied by the Commission's comments, to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions <u>prior to presenting a proposal for the subsequent FP</u> (Article 5(2)) 	<ul style="list-style-type: none"> to <u>communicate</u> the conclusions of this assessment, accompanied by the Commission's comments, to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions (Article 6(2)) in good time and <u>before submitting its proposal for the subsequent FP</u> (recital (20)) 	
Final evaluation			
(see Specific Programme Decision)			<ul style="list-style-type: none"> to carry out an <u>external evaluation two years following the completion of FP7 (i.e. 2015)</u> by independent experts of its rationale, implementation and achievements (Article 7(3)) to <u>communicate</u> the conclusions thereof, accompanied by its observations, to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (Article 7(3))
Independent experts			
	<ul style="list-style-type: none"> to ensure in the <u>selection of the independent experts</u> that account is taken, in a balanced manner, of different scientific, industrial and user communities and to publish the names of experts and their individual qualifications following their appointment (Article 5(3)) <p>(for additional provisions regarding independent experts see Council Decision 1999/65/EC, Article 9)</p>	<ul style="list-style-type: none"> the Commission shall designate independent experts to assist with the evaluation required under the Sixth Framework Programme and the specific programmes, and also for the assistance referred to in Article 10(6) and the second subparagraph of Article 18(1) it may in addition set up groups of independent experts to advise on the implementation of Community research policy (for provisions regarding independent experts see Regulation (EC) No 2321/2002, Article 11) 	
Example of one specific programme decision — Council			
Council Decision 94/801/EC of 23 November 1994: Article 4	Council Decision 1999/168/EC of 25 January 1999: Article 4 and 6	Council Decision 2002/834/EC and 2002/835/EC of 30 September 2002 — Article 8	Council Decision 2006/971/EC of 19 December 2006 concerning the specific programme 'Cooperation': Article 8

FP4 (1995 to 1998) ⁽¹⁾	FP5 (1999 to 2002) ⁽²⁾	FP6 (2003 to 2006) ⁽³⁾	FP7 (2007 to 2013) ⁽⁴⁾
Monitoring			
<ul style="list-style-type: none"> to continually and systematically monitor, with appropriate assistance from independent external experts, the progress within the programme (Article 4(1)) 	<ul style="list-style-type: none"> to monitor, with appropriate assistance from independent external experts the implementation of the specific programme and, where appropriate, to submit proposals to Council for adapting it, in accordance with Article 5(1) of FP5 Decision (Article 4(a)) 	<ul style="list-style-type: none"> to arrange for the independent monitoring and assessment provided for in Article 6 of the FP6 to be conducted concerning the activities covered by the specific programme (Article 8(2)) 	<ul style="list-style-type: none"> to arrange for the independent monitoring, assessment and review provided for in Article 7 of the Framework Programme to be conducted concerning the activities carried out in the fields covered by the specific programme (Article 9) monitoring information to be provided to the Programme Committee according to Article 8(4) by Commission is specified in Annex V.
Article 173 — Annual report			
		<ul style="list-style-type: none"> to regularly report on the implementation of the specific programme and, including information on financial aspects and the use of instruments, in accordance with Article 4 of FP6 Decision (Article 8(1)) 	<ul style="list-style-type: none"> to include main findings in respect [of the Risk-sharing finance facility] to the annual report on RTD activities which it will send to the European Parliament and the Council pursuant to Article 173 of the Treaty (Annex III).
Five-year assessment (5YA)			
<ul style="list-style-type: none"> to have an external assessment conducted by independent qualified experts of the activities carried out within the domains covered by the programme and their management during the five years preceding this assessment (Article 4(2)) 	<ul style="list-style-type: none"> to arrange for the external assessment provided for in Article 5.2 of the FP5 Decision to be conducted concerning the activities carried out in the fields covered by the specific programme (Article 4(b)) 		
Final evaluation			
<ul style="list-style-type: none"> to provide, on completion of each specific programme, an independent final evaluation of the results achieved compared with objectives (Article 4(3)) to forward this evaluation to the European Parliament, the Council and the Economic and Social Committee (Article 4(3)) 			
Independent experts			
	<ul style="list-style-type: none"> to draw-up the terms of reference for the external assessment provided for in Article 5(2) of FP5 Decision (Article 6(2)) 		

FP4 (1995 to 1998) ⁽¹⁾	FP5 (1999 to 2002) ⁽²⁾	FP6 (2003 to 2006) ⁽³⁾	FP7 (2007 to 2013) ⁽⁴⁾
Basic information on FP structure			
— FP4 (EC): 7 specific programmes comprising 17 sub-programmes plus two clusters of activities	— FP5 (EC): 7 specific programmes, of which 4 are thematic programmes and 3 are horizontal programmes	— FP6 (EC): 20 programmes, of which 10 are regrouped under 7 thematic priorities, and 4 programmes which relate to cross-cutting research activities, in the Specific programme 'Focusing and integrating Community research'. The Specific programme 'Structuring the European Research Area' represent 4 programmes, and the Specific programme 'Strengthening the foundations of the European Research Area' 2 specific programmes	— FP7 (EC): 24 specific sub-programmes, of which 10 are regrouped in the specific programme 'Cooperation', one relates to the specific programme 'Ideas', 5 sub-programmes relate to the specific programme 'People' and 8 sub-programmes relate to the specific programme 'Capacities' (one of which is devoted to the non-nuclear activities of the JRC)
— Total budget (EC): 11 046 million euro (i.e. 2 761 million euro p.a.)	— Total budget (EC): 13 700 million euro (i.e. 3 425 million euro p.a.; + 24 %)	— Total budget (EC): 16 270 million euro (i.e. 4 061 million euro p.a.; + 19 %) after EU-25 enlargement increased to — Total budget (EC): 17 883 million euro (i.e. 4 447 million euro p.a.; + 10 %)	— Total budget (EC): 50 521 million euro (i.e. 7 217 million euro p.a.; + 62 %)
— 7 Directorates-General: Research, Information Society, Energy, Transport, Fisheries, Agriculture, Enterprise	— 5 Directorates-General: RTD, Information Society, Energy and Transport, Fisheries, Enterprise	— 4 Directorates-General: RTD, Information Society, Energy and Transport, Enterprise	— 5 Directorates-General: RTD, Information Society and Media, Energy and Transport, Enterprise and Industry, Joint Research Centre-(JRC); in addition 'European Research Council' (ERC) and executive agencies

⁽¹⁾ **FP4:** Information and Communications Technologies (Advanced Communications Technologies and Services (ACTS); Information Technologies (ESPRIT); Telematics Applications); Industrial Technologies (Industrial and Materials Technologies (BRITE/EURAM); Standards, Measurements and Testing (SMT)); Environment (Environment and Climate, Marine Sciences and Technologies (MAST III)); Life Sciences and Technologies (Biotechnology BIOTECH 2), Biomedicine and Health (BIOMED 2), Agriculture and Fisheries (FAIR), Ethical, Legal and Social Aspects (ELSA)); Energy (Research and Development (JOULE), Demonstration (THERMIE)); Transport (Transport Research Programme), Targeted Socio-Economic Research (Targeted Socio-Economic Research (TSER), European Technology Assessment Network (ETAN)); Cooperation with Third Countries and International Organisations (International Cooperation (INCO)); Dissemination and Exploitation of Results (Innovation Programme); Stimulation of the training and mobility of researchers (Training and Mobility of Researchers (TMR), Marie Curie Fellowship Association).

⁽²⁾ **FP5:** 4 Thematic Programmes (Quality of Life and management of living resources (Quality of Life), User-friendly information society (IST), Competitive and sustainable growth (GROWTH), Energy, environment and sustainable development (EESD)); 3 Horizontal Programmes (Confirming the international role of Community research (INCO 2), Promotion of innovation and encouragement of SME participation (Innovation/SMEs), Improving the human research potential and the socio-economic knowledge base (Improving)).

⁽³⁾ **FP6:** 'Focusing and integrating Community research'

- 7 thematic priorities representing 10 programmes: Life sciences, genomics and biotechnology for health (incl. Advanced genomics and its applications for health, Combating major diseases), Information society technologies, Nanotechnologies and nanosciences, knowledge-based multifunctional materials and new production processes and devices, Aeronautics and space, Food quality and safety, Sustainable development, global change and ecosystems (incl. Sustainable energy systems, Sustainable surface transport, Global change and ecosystems), Citizens and governance in a knowledge-based society,
- cross-cutting research activities covering a wider field of research representing 4 programmes: Policy support and anticipating scientific and technological needs, Horizontal research activities involving SMEs, Specific measures in support of international cooperation; Non-nuclear activities of the Joint Research Centre; 'Structuring the European Research Area' representing 4 programmes: Research and innovation, Human resources, Research infrastructures, Science and society 'Strengthening the foundations of the European Research Area' representing 2 programmes: Support for the coordination of activities, Support for the coherent development of policies.

⁽⁴⁾ **FP7:** specific programme 'Cooperation' representing 10 sub-programmes (Health; Food, agriculture and biotechnology; Information and Communication technologies; Nanosciences, nanotechnologies, materials and new production technologies; Energy; Environment (including climate change); Transport (including aeronautics); Socio-economic sciences and the humanities; Security; Space), specific programme 'Ideas' representing one sub-programme (European Research Council), specific programme 'People' representing 5 sub-programmes (Initial training; Life-long training; Industry-academia; International dimension; Specific actions), specific programme 'Capacities' representing 8 sub-programmes (Research infrastructures; Research for the benefit of SMEs; Regions of knowledge; Research potential; Science in society; Coherent development of research policies; International co-operation; Non-nuclear activities of the JRC).

ANNEX III

MAIN DIFFERENCES IN EX-POST EVALUATION OF RESEARCH PROGRAMMES BETWEEN EU (FP4 — FP6), USA AND CANADA

	EU (FP4 — FP6)	USA	Canada
Systemic issues			
Intervention logic	not explicit	mostly in place, implicitly required by PART (*)	required by RMAF (**) as part of programme proposal
Evaluation strategy	not comprehensive	required by PART (*)	required by RMAF (**) as part of programme proposal
Specifying expected results	no	required by PART (*) (incl. use of quantifiable indicators)	required by RMAF (**) as part of programme proposal
Evaluation questions asked by	Commission as programme manager	agency as programme manager (assessed within OMB framework)	agency as programme manager (assessed within TBS framework)
Organisational issues			
Design phase	during programming period	during programming period	completed before programme starts
Use of panels	'Ad hoc' panels for specific evaluation exercises	Depending on agency/programme	Permanent panels as standing committees
Periodicity	fixed dates (linked to programming period, covering preceding 5 years)	2-3 year cycle (linked to budget approval)	upon programme renewal

Notes:

(*) Programme Assessment Rating Tool (PART) since 2002: see Office of Management and Budget (OMB), Executive Office of the President, USA.

(**) Results-based Management and Accountability Framework (RMAF): see Treasury Board of Canada Secretariat (TBS).

ANNEX IV
POSSIBLE PANEL STRUCTURE FOR FP7

Overall FP evaluation panel

Cooperation	Ideas	People	Capacities
Total budget: 32 413 million euro	Total budget: 7 510 million euro	Total budget: 4750 million euro	Total budget: 5 848 million euro
10 sub-programmes: <ul style="list-style-type: none"> — Health (6 100 million euro) — Food, agriculture and biotechnology (1 935 million euro) — Information and Communication technologies (9 050 million euro) — Nanosciences, nanotechnologies, materials and new production technologies (3 475 million euro) — Energy (2 350 million euro) — Environment (including climate change) (1 890 million euro) — Transport (including aeronautics) (4 160 million euro) — Socio-economic sciences and the humanities (623 million euro) — Security (1 400 million euro) — Space (1 430 million euro) 	1 sub-programme: <ul style="list-style-type: none"> — European Research Council <p><i>(programme implementation by Executive agency)</i></p>	5 sub-programmes: <ul style="list-style-type: none"> — Initial training — Life-long training — Industry-academia — International dimension — Specific actions <p><i>(programme implementation by Executive agency)</i></p>	8 sub-programmes: <ul style="list-style-type: none"> — Research infrastructures (1 715 million euro) — Research for the benefit of SMEs (programme implementation by Executive agency) (1 336 million euro) — Regions of knowledge (126 million euro) — Research potential (340 million euro) — Science in society (330 million euro) — Coherent development of research policies (70 million euro) — International co-operation (180 million euro) — non-nuclear activities of the JRC (1 751 million euro)
10 sub-panels	1 sub-panel	1 sub-panel	1 sub-panel

Note: Panel structure as recommended by ECA.
Source: FP7 Decision.

THE COMMISSION'S REPLIES

I. The Community RTD Framework Programmes (FPs) have developed over the last decades into a wide-ranging, but coherent system of support mechanisms for researchers in the 27 EU Member States, the Associated Countries and beyond.

The FPs are today the biggest public multinational research funding system worldwide.

The FP currently represents around 5 % of total EU public funding of research. In terms of annual Community financial participation FP7 will be some 60 % larger than FP6.

Evaluating such a unique scheme is a particular challenge for all actors involved. Evaluation and monitoring arrangements have gone through significant changes over recent years, most recently with the launch of FP7 in 2007.

The evaluation and monitoring system in place at the Commission is based on the direct responsibility of the implementing Commission services. This is where the highest degree of technical expertise is available.

The evaluations at the Framework Programme level are carried out through coordinated efforts across all Directorates General (DGs).

IV. The Commission welcomes the Court's observation that the DGs implementing the FPs (research DGs) have complied with the Commission requirements, including those specified in the 2000 and 2002 Communications on evaluation.

Significant changes have been introduced to FP evaluation, including a major restructuring for FP7.

The Commission also considers that:

- (i) the FPs have always been based on a sound intervention logic. In FP7 this is made more explicit, which will further strengthen the Commission's evaluation and monitoring system;
- (ii) the main strategic elements of the Commission's approach to evaluation and monitoring were defined in the successive legislative acts for the FPs since the mid-1990's.

Since the Commission adopted its last Communication on research evaluation in 1996, a process of constant refinement of the Commission's approach to FP evaluation occurred, and significant changes to the evaluation exercises were made.

During the period audited by the Court, major evaluation exercises of the FPs, the so-called 'Five-year assessments', were carried out every four years;

- (iii) the FPs are jointly implemented by the 'research DGs', which act independently and cooperate where necessary. The Commission's central services, and DG Budget in particular, participate actively in the different steps related to the evaluation process.

Given the Commission's decentralised organisation and the specificities of each scientific field, evaluations need to be managed by the operational services in charge.

The Commission-internal 'RTD evaluation Network' has proved to be very useful for coordination.

The Commission agrees that panels should play an important role in evaluation and welcomes the audit finding that its panels for the 'Five-year assessments' and other major evaluations were perceived by institutional stakeholders as a guarantee for the independence of the approach followed by the Commission during the period covered by the audit;

- (iv) providing direct methodological guidance to evaluations is constrained by the need to maintain the independence of the evaluators. The Commission acknowledges past difficulties with data gathering and accepts that less is known about longer term outcomes and impacts of FPs, as is true of most evaluations of public interventions in the research area and beyond;

- (v) the timing of most evaluations is specified in the FP legislation adopted by the European Parliament and the Council. Overall, the Commission considers its evaluation and monitoring system of the FPs to be effective. This view was shared by the European Parliament and the Council when adopting the Commission's recent legislative proposals for FP7 in 2006.

V. In view of the Court's recommendations, the Commission:

- (i) points to the more explicit intervention logic embodied in the in FP7 legislation (adopted by the European Parliament and the Council in 2006) and will reinforce the use of an explicit intervention logic in its future legislative proposals. Moreover, the Commission intends to further clarify its operational objectives for FP7 at the level of work programmes and to further develop its approach to performance measurement;
- (ii) will further develop its evaluation strategy for FP7, based on the proposals in the FP7 Impact Assessment. Care will be taken to ensure an appropriate level of flexibility so that each operational service can determine its specific evaluation needs;
- (iii) while it does not consider it necessary to set up a joint evaluation office, agrees to consider other options for improving coordination.

Panels should be convened at the earliest possible point to support a specific evaluation. Dedicated panels, convened for specific purposes, are however more likely to provide benefits in terms of clarity and focus;

- (iv) has already taken action to better exploit existing and new data sources for its evaluations.

FP7 reporting requirements were revised partly on the basis of evaluation and monitoring needs.

In addition, the Commission is currently developing a new IT system for FP7 which is intended to provide a more robust basis in terms of data on projects.

The Commission will consider the possibility for a new manual on FP evaluation, taking account of recently produced examples;

- (v) the Commission intends to develop and strengthen its forward planning of evaluations under FP7, reflecting the possible advantages for consistency, and take-up of findings.

1-13. The Community RTD Framework Programmes (FPs) have developed over the last decades into a wide-ranging, but coherent system of support mechanisms for researchers in the 27 EU Member States, the Associated Countries and beyond.

14. The DGs implementing the RTD FPs (research DGs) have formally complied with the Commission's requirements for evaluation, including those specified by the 2000 and 2002 Communications on evaluation (see paragraph 54 of the Court's text).

The FPs are today the biggest multinational research funding system worldwide (see paragraph 6 of the Court's text). In terms of annual Community financial participation FP7 will be some 60 % larger than FP6 (see figure 1 of the Court's text).

22. The Commission agrees with the importance of a sound intervention logic for the programme design and share the view that the mandate specified in Article 163 of the Treaty is a key element in this respect. In addition, the contribution of RTD activities funded by the Commission to other Community policies and objectives is of great importance. More recently, the Commission's view of the role of EU RTD investment was presented in detail in both the renewed 'Lisbon' and 'Sustainable Development' strategies ⁽¹⁾.

Evaluating such a unique scheme is a particular challenge for all actors involved. Evaluation modalities have gone through significant changes over recent years (see figure 2 of the Court's text), most recently with the launch of FP7 in 2007 (See for example the replies to paragraphs 23, 30, 41, 46, 72, 76-78, 93).

⁽¹⁾ Communication from the Commission to the Spring European Council — Implementing the renewed Lisbon Strategy for Growth and Jobs (COM(2006) 816).

23. The FPs have always been based on a sound intervention logic. In FP7 this is made more explicit ⁽¹⁾, which will further strengthen the Commission's evaluation and monitoring system.

25. A number of countries including some Member States are developing this approach, but this is by no means a uniform approach. Lessons from the USA and Canada may be less easily transferable than from EU Member States.

26. The Commission considers that it has had in place an effective evaluation and monitoring system during the period audited by the Court (see also reply to paragraph 23).

28. It is challenging, in particular given the uncertainty of research, to ensure all objectives fit precisely within the strict definitions of SMART criteria. All work programmes under FP7 contain expected impact statements, which correspond with the objectives, rationale and activities set out in the Specific Programmes.

29. Greater attention has been given to shorter-term achievements during the period audited by the Court; there is room for more long-term analyses.

30. At the FP level a monitoring system is currently being developed by the Commission for FP7 aiming to provide a more coherent and systematic use of performance indicators. Several separate initiatives for the development of performance indicators are also ongoing within the different services implementing FP7.

34. The Commission is already committed to ensuring greater clarity in programme intervention logic, building on the advances embodied in the FP7 legislation and Impact Assessment and it will consider the Court's recommendation for future legislative proposals.

35. The Commission welcomes the Court's recommendation to further clarify FP7 programme objectives, and, within the constraints of the relevant basic acts adopted by the European Parliament and the Council, will work to achieve this.

36-37. The Commission welcomes the Court's recommendations and is taking initiatives for the further development of performance measurement and for the modification of the ABM/ABB nomenclature.

(1) For instance, the Commission proposal for FP7 was developed on the basis of a detailed *ex-ante* Impact Assessment, which included economic modelling of the potential impacts of the research activity on issues such as economic growth and job creation (see European Commission, Commission Staff Working Paper, 'Impact Assessment and Ex-Ante Evaluation for the proposals for the Council and European Parliament decision on the 7th Framework Programme (EC and Euratom)', SEC(2005) 430). The FP7 Decisions provide a clear structure which links objectives, rationale and activities. In addition, the FP7 Specific Programme Decisions set out objectives, approach and activities.

38-39. The Commission considers that its evaluation strategy for the FPs should set out common principles for the organisation and development of FP evaluation and monitoring. The strategy should provide a statement of purpose, define key components of the system including major exercises and elaborate a vision for the future. It should remain focused at the strategic level in view of the Commission's decentralised evaluation system for the FPs.

41. The main strategic elements of the Commission's approach to evaluation and monitoring were defined in the successive basic acts for the FPs since the mid-1990s (see *Annex II*).

Since the Commission adopted its Communication on research evaluation in 1996, a process of constant refinement of its approach to FP evaluation occurred and significant changes to the evaluation exercises were made (see figure 2 of the Court's text).

For the current programming period, the Commission's planned evaluation strategy is set out more explicitly in the FP7 *ex-ante* Impact Assessment (SEC(2005) 430). It covers the design of the evaluation system, how it will be implemented and a detailed timetable of planned FP level evaluations.

More work will be necessary to translate this strategy into specific plans and the Commission will use specific coordination structures for this purpose.

43. While acknowledging the Court's observation, the Commission notes that the situation in the USA and Canada is considerably more nuanced than could be presented by single evaluation strategies.

44. Differences in the use of and approach to evaluation across the FP reflect the decentralised Commission system, which links evaluation with decision making.

46. Given the Commission's decentralised organisation, the question of the appropriate level for evaluation activity needs to be answered by the operational services in charge (see also replies to paragraph 41 and 49).

— (second indent) The 2004 Five-Year Assessment (1999-2003), coordinated by DG Research, provided an evaluation at the overall FP level. The decision of whether to carry out separate exercises at programme levels reflected the needs of the services involved and the different contexts and scale of activities concerned.

47. The 2004 Five-Year Assessment provided a coordinated, comprehensive, on-time evaluation at the overall FP level, in compliance with the Financial Regulation. This was carried out by distinguished external experts with the results widely disseminated.

48. See replies to paragraphs 41 and 46.

49. For the current programming period, the Commission's planned evaluation strategy is set out in the FP7 *ex-ante* Impact Assessment (SEC(2005) 430). It covers the design of the evaluation system, how it will be implemented and a detailed timetable of planned FP level evaluations.

The Commission accepts that more work is necessary to translate this strategy into specific plans and will use specific coordination structures for the purpose.

The Commission considers however that a minimum level of detailed evaluation cannot always be set due to the need to retain sufficient flexibility for management to determine the appropriate evaluation needs.

51. FPs are jointly implemented by the 'research DGs', which act independently and cooperate where necessary.

52. EuropeAid has operational and management responsibilities that it exercises in collaboration with Commission delegations and implementing bodies in third countries. This makes it difficult to consider this part of the Commission as an example as it is not comparable with the decentralised system of responsibilities and roles in the research area.

53. The Commission's internal 'RTD evaluation network' provides informal means for the development of an evaluation strategy and for coordinating FP level evaluation exercises. This flexible mechanism has proved to be very useful.

55. The recent revision of the Commission's Internal Control Standards, adopted on 16 October 2007 (SEC(2007) 1341) encompasses the 21 evaluation standards and compliance with these standards will therefore be included in the scope of the DGs' annual activity reports.

The Commission's central services, DG Budget in particular, participate actively in the different steps related to the evaluation process.

The evaluation activities of the services may be subjected to audit by the internal audit capability and the internal audit service.

In the future, the assessment of minimum quality requirements will use the recently developed Commission's 'Quality Assessment Framework' ⁽¹⁾. Since 2007, the quality control for prospective Impact Assessment has been assisted by the Commission's 'Impact Assessment Board'.

⁽¹⁾ The 'Quality Assessment Framework' is available on http://ec.europa.eu/budget/documents/evaluation_en.htm.

56. The comparison provided in *Annex III* covers the period audited by the Court. Therefore, it does not take account of significant improvements under FP7, notably as regards a new evaluation strategy, a more explicit intervention logic and clear statements of expected results in the work programmes.

The Court highlights certain Member States where the responsibility for evaluation and funding are separate. Equally, other examples exist (such as the UK) where evaluation is part of programme management (as is the case in the Commission). On the key issue of independence, the Court's audit has shown that external stakeholders are satisfied with the Commission's performance on this point (see paragraph 58 and figure 4 of the Court's text).

59.

— (first indent) Ensuring an adequate involvement of high-level members for panels convened over long periods is difficult.

— (second indent) Terms of reference for FP level panels are set by Commission Decision, following extensive informal stakeholder consultation, including the formal opinion of the FP7 Programme Committees made up of representatives of the Member States and associated countries. This will also be the case for the forthcoming FP6 *ex-post* evaluation in 2008.

62. The target of 0,5 % of a programme's budget to be devoted to evaluation is recommended by the Commission as an average. The resources actually allocated depend much on the nature of a programme.

66. While the Commission does not consider it necessary to set up a joint evaluation office, it agrees to consider other options for improving coordination.

The link between evaluation and management decision making can best be achieved by evaluation functions situated in each DG and by giving a new impetus to the Commission's existing internal 'RTD evaluation network' with the support of central services. This is also in line with the legislative and budgetary processes, as required by the Financial Regulation, the 2007 Commission Communication on evaluation and the procedures in the context of ABM.

67. The Commission agrees that panels can play an important role in evaluation. Whilst it also agrees that panels should be convened at the earliest possible point to support a specific evaluation, it does not consider it appropriate to establish standing panels throughout the programming period.

A dedicated panel, convened for a specific purpose, is more likely to provide benefits in terms of clarity and focus.

68. The Commission will carefully assess the needs for resources to carry out the evaluation of the FPs in the forthcoming years.

70-71. Methodological guidance for evaluation, including available tools and glossary of terms, is provided by the DG Budget guide ⁽¹⁾ which services, including the 'research DGs', are currently updating.

In addition, specific evaluation guidance was developed by DG Enterprise in 2002 covering all activities for which this DG is responsible, including the FPs.

Further supporting initiatives will be considered for the re-launched 'European RTD Evaluation Network' (composed of experts from the Member States and associated countries) and the Commission's internal 'RTD Evaluation Network'. The Commission will also continue to support independent evaluation methodology studies and workshops.

72. The Commission agrees that the process for the routine collection of data from projects has only been partially effective. Project reporting procedures were revised under FP6, and will be further improved for FP7 (e.g. periodic reporting will be simplified; interim questionnaires have been suppressed; final reporting will become more useful, notably with regard to the collection of data on project implementation and results).

74. The 'research DGs' are currently developing a new IT system for FP7 with the support of the Informatics DG. Such an enhanced IT support is intended to provide a more robust basis for evaluation and monitoring.

76-78. The sourcing of evaluation data was only partially effective during the period covered by the audit. The Commission is introducing improvements under FP7 (see reply to paragraph 72).

79. The Commission aims to encourage the types of quantitative studies proposed by the Court.

81. In the Commission's view, many methodologically more interesting evaluations were carried out recently (see paragraph 84) and a distinction should be made between the large-scale evaluations and more routine exercises.

82. The Commission accepts that in general, evaluation work has given greater attention to shorter-term achievements and agrees there is room for more analysis with regard to outcomes and impacts. However, such an approach requires a longer term perspective and therefore the Commission's evaluations are typically comparable to similar work elsewhere.

⁽¹⁾ European Commission, DG Budget, 'Evaluating EU Activities — A Practical Guide for the Commission services', July 2004, is available at http://ec.europa.eu/budget/publications/other_en.htm.

83. The Commission seeks to respect the independence of evaluation experts, notably on the choice of the appropriate methodologies.

For the forthcoming *ex-post* evaluation of FP6 in 2008 it is proposed to carry out studies on behavioural additionality, networks analysis and the use of scientometric approaches.

84. A meta-analysis of previous evaluation studies was carried out by an independent expert for the 2004 Five-Year Assessment (1999-2003) ⁽²⁾.

— (second indent) In addition to the studies referred to by the Court, further major studies have been launched in 2007 ⁽³⁾.

85. The Commission agrees that there is a further need to develop new methodological approaches to address the changing nature of the FPs. Such a methodological development is, and will continue to be, supported by the Commission.

86. As the questions listed in the Terms of Reference for the 2004 Five-Year Assessment (1999-2003) were indicative only, the panel was free to choose the areas on which to focus and to determine its working modalities. Panel members have commented on the excellence of the exercise ⁽⁴⁾.

87. The very nature of a research programme, dealing with uncertainty, means there are limits to the extent to which quantifiable objectives can be specified.

90. In 2007, the Commission re-launched its 'European RTD Evaluation Network', comprised of Member States officials and experts in evaluation, which it plans to use to promote more coherence in EU research evaluation activities, notably with regard to the impact studies at Member State level.

91. The Commission will consider the possibility for a new manual on FP evaluation, taking account of recently produced examples ⁽⁵⁾.

⁽²⁾ E. Arnold, 'What the Evaluation Record tells us about the Framework Programme Performance', 2004.

⁽³⁾ DG Research, 'Structuring Effects of Community Research — The Impact of the Framework Programme on Research & Technological Development (RTD) Network Formation', 2007; DG Information Society and Media 'Effectiveness of IST-RTD Impacts on the EU Innovation System', 2007.

⁽⁴⁾ See C. Mandl, 'Evaluating Evaluations or The Case for Action Research', Austrian Platform Newsletter, No 30, June 2007.

⁽⁵⁾ See for instance European Commission, Joint Research Centre, 'Assessing the Socio-Economic Impact of RTD-Policies' (EPUB), 2002, and University of Manchester, 'Assessing the Socio-economic Impacts of the Framework Programme' (ASIF), 2002, both available at http://circa.europa.eu/Public/irc/rtd/fiveyearasskb/library?l=/iii-knowledge_base/040-evaluation_methodolo&vm=detailed&sb=Title. These reports have been widely distributed and were made available to contractors and experts carrying out evaluation studies on behalf of the Commission.

92. Project reporting procedures were revised under FP6, and will be further improved for FP7 (see reply to paragraph 72).

93. At the FP level, a monitoring system is currently being developed by the Commission for FP7 aiming to provide a more coherent and systematic use of performance indicators (see reply to paragraph 30).

The 'research DGs' are currently developing a new IT system for FP7 with the support of the Informatics DG. Such an enhanced IT support is intended to provide a more robust basis for evaluation and monitoring.

94. The Commission shares the Court's concerns as regards the problem of the reporting burden and has taken this specifically into account when specifying the reporting requirements for FP7 (see reply to paragraph 72).

95. The Commission is currently engaged in efforts to exploit existing and new data sources for evaluation and monitoring (see reply to paragraph 72).

96. Within FP7, there is a specific action line under theme 8 (Social sciences and humanities) to address innovative approaches to evaluation and monitoring.

97. The Commission has re-launched in 2007 its European RTD Evaluation Network (see replies to paragraphs 70-71 and 90), which it plans to use to promote more coherence in EU research evaluation activities, notably the impact studies of Member States.

The Commission also intends to maintain its position in the vanguard of support to international networking and sharing of research evaluation practice through workshops and other events.

99. Evaluation timing is determined by the requirements specified in the basic acts adopted by the European Parliament and the Council. The Commission accepts that less is known about long term impacts and agrees there is room for more work in this area.

100. The Commission agrees that the Five-Year Assessments tended to focus more on shorter term assessment than on the longer term. This is explained by factors including the interest of external stakeholders, such as the European Parliament and the Council, on quick feedback on FP performance, the challenge of attributing impacts to research over the longer-term and the focus chosen by the independent panels carrying out the exercises. The Commission accepts however that more attention needs to be given to assessing longer-term effects.

108. The Commission is in the process of planning the main evaluation activities under FP7 and will take the Court's recommendation carefully into consideration. The Commission accepts that less is known about long term impacts and agrees there is room for more work in this area.

109. As noted by the Court at paragraph 107, the Commission provides a summary of all findings and actions taken for research evaluations in its Annual Evaluation Review. The Commission is committed to full transparency for its evaluation work and will take up the recommendation of the Court to review the means for communicating and disseminating findings.

110-111. The Commission agrees with the Court that evaluation of research, notably for a programme as comprehensive as the FP, is very challenging.

It notes that very significant improvements have been made to the system over the last 25 years and that the system overall has been effective for policy makers, research managers and other stakeholders.

112. The Commission welcomes the Court's observation that the DGs implementing the FPs have complied with the Commission requirements, including those specified in the 2000 and 2002 Communications on evaluation.

Significant changes have been introduced to FP evaluation, including a major restructuring for FP7, as set out in the strategy provided by the Impact Assessment (SEC(2005) 430).

113. Concerning the relevance and usefulness of evaluations, the Commission can point to the generally positive feedback it has received, notably with recent key exercises such as the 2004 Five-Year Assessment (1999-2003). There was extensive dissemination of the findings of the Five-Year Assessment including presentations to key bodies such as the European Parliament, the Council, CREST and the FP Programme Committee. Given that it was a mandatory requirement that this exercise be carried out before the Commission's proposal for a new Framework Programme (in this case FP7), its overall acceptability was indicated by the eventual adoption of FP7.

This is not to say that improvements cannot be made and the Commission accepts fully the need to make carefully considered changes which build on the elements of the current approach which already work well.

114. The Commission welcomes the Court's recognition of the importance of evaluation and monitoring and, building on the significant improvements already made, will continue further updating its current practices in the light of new political and practical realities.

115. The Commission is already committed to ensuring greater clarity in programme intervention logic, building on the advances embodied in the FP7 legislation and Impact Assessments, and it will consider the Court's recommendation for future legislative proposals.

116. Building on the significant improvements already made to FP evaluation, the Commission agrees that there is still room for improvement including closer co-ordination at operational level among the services concerned. The Commission will continue its efforts to support developments in evaluation methodology.

117. As regards the organisational arrangements, the Commission considers that:

- (first indent) panels should be convened at the earliest possible point to support a specific evaluation. Dedicated panels, convened for specific purposes, are however more likely to provide benefits in terms of clarity and focus,
 - (second indent) the coordination between the DGs implementing the FPs needs to be reinforced. In addition to the new FP7 evaluation strategy, this will be achieved by using specific coordination structures to complement the Commission-internal 'RTD evaluation network', rather than setting up a joint evaluation office.
-