



European Union
Regional Policy



Developing Evaluation Capacity

A study for the Commission

February 2008



EUROPOS SOCIALINIAI, TEISINIAI IR EKONOMINIAI PROJEKTAI

**FINAL REPORT
ON THE FRAMEWORK
TO ANALYSE THE DEVELOPMENT OF EVALUATION
CAPACITY IN THE EU MEMBER STATES**

2nd edition

Contract No.2006.CE16.0.AT.023

11 September 2007

CONTENTS

ABBREVIATIONS.....	4
EXECUTIVE SUMMARY	5
1. INTRODUCTION	14
1.1. BACKGROUND.....	14
1.2. OBJECTIVES OF THE STUDY.....	15
1.3. STRUCTURE OF THE REPORT	16
2. RESEARCH METHODOLOGY	17
2.1. CONCEPTUALISING EVALUATION CAPACITY AND DEVELOPING BENCHMARKING FRAMEWORK	17
2.2. APPLICATION OF THE BENCHMARKING FRAMEWORK	17
2.2.1. Survey.....	18
2.2.2. Country Case Studies	19
2.3. REVISION OF THE BENCHMARKING FRAMEWORK	21
2.4. DEVELOPING RECOMMENDATIONS	21
3. EVALUATION CAPACITY DEVELOPMENT LITERATURE REVIEW	22
3.1. DEFINITION OF EVALUATION CAPACITY AND ITS RATIONALE	22
3.1.1. Definition of Evaluation Capacity.....	22
3.1.2. Rationale: Why Evaluation Capacity Development Matters.....	23
3.2. KEY DIMENSIONS OF EVALUATION CAPACITY.....	24
3.2.1. Introduction: The Basic Evaluation Capacity Development Framework.....	24
3.2.2. Evaluation Demand.....	26
3.2.3. Evaluation Supply	26
3.2.4. Evaluation Architecture	26
3.2.5. Institutionalisation	27
3.2.6. Summary and Conclusion.....	28
3.3. EVALUATION CAPACITY CRITERIA AND INDICATORS	28
3.3.1. Introduction.....	28
3.3.2. Furubo and Sandahl Benchmarking Criteria.....	28
3.3.3. EVALSED Evaluation Capacity Indicators.....	29
3.3.4. World Bank Performance Indicators for Evaluation Capacity Building	31
3.3.5. Summary and Conclusion.....	31
3.4. CRITICAL SUCCESS FACTORS FOR EVALUATION CAPACITY DEVELOPMENT	32
3.5. LITERATURE REVIEW: SUMMARY AND CONCLUSIONS	35
4. EVALUATION CAPACITY DEVELOPMENT BENCHMARKING FRAMEWORK	36
4.1. CONTEXT FOR EVALUATION OF EU STRUCTURAL AND COHESION FUNDS.....	36
4.1.1. Some Features of the EU Structural and Cohesion Funds Evaluation Model	36
4.1.2. Future Challenges.....	37
4.2. WORKING DEFINITION OF EVALUATION CAPACITY	37
4.3. STRUCTURE OF EVALUATION CAPACITY BENCHMARKING FRAMEWORK	38
4.3.1. Demand Dimension.....	39
4.3.2. Supply Dimension	40
4.3.3. Dissemination and Utilisation Dimension	40
4.3.4. Institutionalisation Dimension.....	41
4.4. EVALUATION CAPACITY DEVELOPMENT BENCHMARKING FRAMEWORK.....	41
5. RESULTS OF THE FIELDWORK.....	47
5.1. STRUCTURAL AND COHESION FUND EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES	47
5.1.1. Survey Statistics	47
5.1.2. The Demand Side Dimension	50
5.1.3. The Supply Side Dimension.....	61
5.1.4. The Dissemination and Utilisation Dimension.....	68
5.1.5. Institutionalisation of Evaluation in the Wider Governance System.....	76
5.2. METHODS FOR APPLICATION OF THE BENCHMARKING FRAMEWORK: LESSONS LEARNT	82
5.2.1. Application of the Benchmarking Framework Through Survey.....	82

5.2.2. <i>Application of the Benchmarking Framework Through Case Study</i>	83
5.2.3. <i>Lessons Learnt</i>	84
5.3. IMPLICATIONS FOR THE BENCHMARKING FRAMEWORK	85
5.3.1. <i>Strengths and Weaknesses of the Benchmarking Framework</i>	85
5.3.2. <i>Revision of Evaluation Capacity Benchmarking Framework</i>	86
6. CONCLUSIONS AND RECOMMENDATIONS	92
6.1. ON THE FRAMEWORK TO ANALYSE EVALUATION CAPACITY DEVELOPMENT	92
6.2. ON EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES	93
BIBLIOGRAPHY	97
ANNEX 1. QUESTIONNAIRES FOR BENCHMARKING EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES	102
QUESTIONNAIRE FOR THE STRUCTURAL AND COHESION FUND ADMINISTRATORS WITH RESPONSIBILITY FOR EVALUATION	102
QUESTIONNAIRE FOR THE EVALUATION MARKET PARTICIPANTS	117
ANNEX 2. CASE STUDY REPORTS	125
REPORT ON ENGLAND CASE STUDY	126
REPORT ON HUNGARY CASE STUDY	135
REPORT ON LITHUANIA CASE STUDY	147
REPORT ON ROMANIA CASE STUDY	162

ABBREVIATIONS

CSF	-	Community Support Framework
DG EMPL	-	Directorate-General for Employment, Social Affairs and Equal Opportunities
DG REGIO	-	Directorate-General Regional Policy
ESF	-	European Social Fund
ERDF	-	European Regional Development Fund
EU	-	European Union
MA	-	Managing Authority
NDP	-	National Development Plan
NGO	-	Non-governmental organisation
NSRF	-	National Strategic Reference Framework
OECD	-	Organisation of Economic Cooperation and Development
OP	-	Operational Programme
SAPARD	-	Special Accession Programme for Agriculture and Rural Development
SPD	-	Single Programming Document
TOR	-	Terms of Reference
UK	-	United Kingdom

EXECUTIVE SUMMARY

Evaluation is a core component of the Structural and Cohesion Fund programming and management processes. Evaluation plays a role in the design of new programmes, in supporting the implementation process and in assessing the effectiveness of interventions. For the 2007-2013 programming period important modifications to the Structural and Cohesion Fund evaluation approach have been introduced aimed at better integrating evaluation into programme management and implementation. Going forward, on-going evaluation will be designed and undertaken in accordance with internal demands at the Member State level. This requires not only strong links between monitoring and evaluation, but adequate evaluation capacity generally to ensure the quality of on-going evaluation and its contribution to decision-making.

The development of evaluation capacity is a multi-dimensional process that requires time, patience and persistence. The literature on the subject suggests that efforts to develop evaluation capacity benefit from a diagnosis and regular monitoring of evaluation systems and practice to detect what is working and areas requiring attention. In addition, experience of other countries and good practice examples can serve as sources of inspiration and learning.

So far systematic and comparable information on Structural and Cohesion Fund evaluation capacity and its development in the EU Member States has been lacking. This research attempted to address this issue by developing a tool to assess and compare evaluation capacity development within Structural and Cohesion Fund administrations. The study was undertaken with the following objectives:

- (1) To conceptualise evaluation capacity and establish a framework of criteria and indicators to analyse its development within the Structural and Cohesion Fund administrations in the EU Member States;
- (2) To test the benchmarking framework by carrying out analysis of the evaluation capacity development within the Structural and Cohesion Fund administrations;
- (3) To present recommendations on how further development of the Structural and Cohesion Fund evaluation capacity can be encouraged.

In order to realise study objectives, the research was carried out in three steps.

Based on a comprehensive literature review, the research started with the analysis and conceptualisation of evaluation capacity. The following working definition of evaluation capacity for the purposes of this project was proposed:

“The ability of Structural and Cohesion Fund administrations to effectively organise and utilise timely, high-quality evaluations of EU-funded programmes”.

The literature review illuminated many of the issues relevant to the development of an evaluation capacity framework. Building on the literature review and the working definition of evaluation capacity, the draft evaluation capacity development benchmarking framework was constructed. It is structured around four core dimensions as follows:

- (1) Demand side dimension which refers to the commissioning side of the evaluation process;
- (2) Supply side dimension which refers to the availability and quality of both external evaluation expertise and socio-economic data resources;
- (3) Dissemination and utilisation of evaluation outputs;
- (4) Institutionalisation dimension which aims to capture whether the wider (i.e., non-EU) administrative environment or culture is conducive to, or enabling of, evaluation (a

secondary, but nonetheless important, focus is on whether EU evaluation requirements influence Member State evaluation practice).

Each evaluation capacity dimension is broken-down into a series of criteria, in turn decomposed into sub-criteria and then broken down into quantitative and qualitative indicators at the lowest level.

Following the development of the draft framework, the next step was to carry out fieldwork aiming to test it and to establish how well it served the purpose of assessing and analysing evaluation capacity development in different EU Member States. A combination of survey and case study methods were used to collect information against the evaluation capacity benchmarking indicators and to test the applicability of the framework.

On the basis of the fieldwork results, strengths and weaknesses of the framework were identified and some revisions of criteria and indicators were proposed. The full framework, revised in the light of the fieldwork, is presented in the table below.

**Framework to Analyse Evaluation Capacity Development:
Dimensions, Criteria, Sub-criteria and Indicators**

Dimension	Criterion	Sub-criteria	Indicators
DEMAND SIDE	Architecture (Integration, location and organisation of Structural and Cohesion Fund evaluation function)	Evaluation Responsibility	1. Location of responsibility for evaluation function within the managing authority/authorities
			2. Existence of dedicated evaluation function (Evaluation Units / national and programme level)
		Coordination	3. Existence of formal mechanisms for evaluation coordination across programmes
		Links with other functions	4. Linkages between evaluation function and other related functions, such as monitoring and programming
	Evaluation Resources	Financial	5. Evaluation budgets (euros)
		Staffing	6. Number of evaluation staff (national and programme level)
			7. Evaluation experience and skills of evaluation staff <ul style="list-style-type: none"> ▪ Evaluation commissioning experience ▪ Evaluation practice experience
		Evaluation flow and coverage	8. Number of evaluations carried out / time period
	Efficiency of evaluation management	Planning	9. Existence of multi-annual and annual evaluation plans at national and programme level
		Terms of Reference	10. Quality of evaluation Terms of Reference (TOR)
		Procurement	11. Average interval between agreement on TOR and appointment of evaluator
			12. Contract award criteria
		Evaluation Steering Committees	13. Existence, composition and functions of evaluation steering committees (national and programme level)
	Quality Assessment	14. Existence of approved set of quality standards	
	Quality of Monitoring Systems	Reliability	15. Quality of indicator data at national /programme level
		Timeliness	16. Average interval between year end and availability of programme indicator data for year in question
SUPPLY SIDE	Availability of socio-economic data	Availability	17. Availability of key socio-economic indicator data (GDP, employment, unemployment, R&D investment) at regional (NUTS II) level
	Availability and quality of evaluation expertise	Availability of evaluation supply on domestic market	18. Number and type of organisations active in the market
			19. Existence of internal evaluation unit capacity
			20. Availability of evaluation training/education options
		Thematic and methodological expertise	21. Breadth and quality of expertise (areas where expertise is particularly strong or deficient: e.g., macroeconomic modelling, horizontal principles, infrastructure, human resources etc.)
	22. Quality of evaluation reports		

Dimension	Criterion	Sub-criteria	Indicators
DISSEMINATION AND UTILISATION	Dissemination	Access to evaluation outputs	23. Public access to evaluation reports
	Utilisation	Procedures for addressing evaluation and follow-up procedures	24. Existence of formal monitoring committee or managing authority responsibilities
			25. Existence of follow-up procedures where evaluation recommendations are agreed
		Impacts on programme design and implementation	26. Degree of influence of ex ante evaluation of 2007-2013 period on: <ul style="list-style-type: none"> ▪ Strategy formulation and programme design ▪ Allocation of resources ▪ Quantification of objectives and indicator selection
		27. Degree of influence of evaluation(s) on ongoing programme management and implementation	
INSTITUTIONALISATION OF EVALUATION IN THE WIDER GOVERNANCE SYSTEM	An enabling governance environment? (Degree of institutionalisation of evaluation in the wider governance system)	Role of Evaluation in State Budget/degree of integration of evaluation into budget process	28. Degree to which evaluations represent routine part of budgetary process
		Role of evaluation in sectoral strategy/policymaking	29. Extent to which evaluations are routinely used in strategy and policy formulation at sectoral/ministry level
		Existence of mandatory evaluation requirements	30. Existence of statutory evaluation requirements
		Existence of an evaluation society or network	31. Membership size
			32. Frequency of meetings/conferences
			33. Key activities
		Influence of EU evaluation requirements	34. Degree of influence of EU Structural Funds evaluation requirements on Member State evaluation practice
Role of civil society and other players	35. Extent to which evaluation findings are drawn on by civil society and social partners		

The fieldwork formed a basis not only for the revision of the benchmarking framework but also provided evidence to support recommendations on further opportunities for evaluation capacity development within Structural and Cohesion Fund administrations.

Conclusions and recommendations which follow from the overall research are presented below. They concern both: (i) methodological issues around the future application and use of the framework; and (ii) substantive issues concerned with evaluation capacity development within Structural Fund administrations.

On the framework to analyse evaluation capacity development

1. ***Limits of benchmarking.*** The initial intention of developing a framework to analyse evaluation capacity development was to produce a benchmarking tool. However, the experience gained demonstrates that setting benchmarks and ranking performance is hardly feasible and appropriate in the field of Structural Fund evaluation, where complex institutional structures, wider administrative environment and historical and cultural contexts need to be taken into account. There is no one best practice approach to Structural and Cohesion Fund evaluation capacity development.
2. ***A flexible analytical framework.*** The research resulted in a tested and ready-to-use analytical framework that systematically addresses multiple aspects of evaluation capacity development and allows for a review of progress in this field. The framework has been informed by a literature review, a consideration of the key features of the Structural and Cohesion Funds evaluation model and the results of the piloting exercise. Going forward, the framework to analyse evaluation capacity development can be used as:
 - (a) a *means for monitoring* and comparing *evaluation capacity development across the EU*, identifying trends, good practices, learning points and areas in need of improvement;
 - (b) a *self-assessment or diagnostic tool by any Member State* willing to undertake an assessment of Structural and Cohesion Fund evaluation capacity at a country/region level. The various criteria and indicators in the framework highlight different aspects of evaluation capacity and allow areas in need of further improvement to be identified.
3. ***Methods of application.*** The most appropriate method to apply the framework is largely through a case study approach with some local input and, on some dimensions, official correspondence.

Recommendations

- For the European Commission***
- Present the framework to analyse evaluation capacity development to Member States and encourage them to use it for (regular) analysis of their own situation (for instance, for evaluation plan purposes).
 - Use the framework to analyse evaluation capacity development in the upcoming ex post evaluation of Structural Fund management systems (namely, concerning evaluation systems).
 - Collect data at regular intervals against indicators for which official communication means can be used (mostly those under the demand side and dissemination and utilisation dimensions).
 - Consider undertaking (commissioning) more in-depth analysis of evaluation capacity development by applying the framework via case studies. In such instances, it is recommended to carry out analyses of different Member States at the same time in order to produce data which is comparable in time and

across the EU.

- Disseminate identified good practice lessons in evaluation capacity development.

For Member States

- Member States are encouraged to apply the framework for their self-assessment. Evaluation capacity analyses are recommended and could be included as projects in their evaluation plans. Such a self-assessment of evaluation capacity is desirable because it can guide the identification of strengths, weaknesses and opportunities for further improvement.

On evaluation capacity development within Structural and Cohesion Fund administrations in the EU Member States

The demand side

1. **Institutionalising evaluation.** The dominant practice in the Member States is to set up formal structures for the evaluation of Structural and Cohesion Funds. Specialised evaluation units operate within managing authorities in a number of Member States. Steering groups and committees to coordinate and manage evaluation issues have been set up in almost all countries. The development of evaluation plans is becoming a prevalent practice in the 2007-2013 programming period; this is a positive feature and should assist Member States in managing evaluation demand in the 2007-2013 period.
2. **Human resources not yet adequate.** The dominant opinion of Structural and Cohesion Fund administrations is that financial allocations for evaluation are generally adequate but that human resources are not, particularly in the EU12. Without formal requirements for evaluation, a sufficient degree of ownership in Member States will be needed to ensure that evaluation is taken seriously. Where institutions are understaffed or where staff are occupied with other functions besides evaluation, it is much more difficult to analyse and plan evaluation needs, which represents a risk factor in the context of the new system of on-going evaluation. This also limits the possibility of time and resources being devoted to the development of specific expertise and skills in evaluation field. In addition, survey evidence suggests that supply of training options is limited in a number of Member States, especially the EU12.
3. **Different levels of awareness.** National authorities (especially at programme level) tend to have different information and different level of awareness about various aspects of Structural and Cohesion Fund evaluation, a finding which further confirms the value of undertaking analysis of evaluation capacity and disseminating its results.
4. **Monitoring systems need further improvement.** The timeliness and quality of monitoring indicator data is still a weak area in a number of countries, especially in the EU12. This can be seen as a risk as well-functioning monitoring systems and the availability of high quality timely data are pre-requisites for good evaluation, especially in the context of on-going evaluation systems.

Recommendations for strengthening the demand side

For the European Commission

- The recommendation which emerges from the interviews with national authorities concerns the value of organising *specific themed workshops at the EU level* which would discuss very practical evaluation issues through a combination of presentations from prominent experts in the field and exchange of Member States' experience. Authorities feel that such issue-specific events are lacking and other training of more general nature cannot fill this gap.

For Member States

- Targeted efforts at strengthening human resources are necessary. *Training* and other evaluation capacity building measures remain crucial if the efficiency of evaluation management is to be increased and if capacity to carry out internal and external evaluations is to be further developed. For this reason, it is recommended to assess, plan and include evaluation capacity building activities in the forthcoming evaluation plans. In addition, information about existing and available training opportunities needs to be better circulated within administrations. This information could be disseminated via institutional websites (sections dedicated to evaluation), intranets, evaluation steering groups, etc.
- Apart from training, periodic *events on exchange of experience* between different national institutions would also contribute to further evaluation capacity development within administration and higher awareness of various aspects of the entire system for evaluation. Such events could focus on such issues as experience gained in commissioning evaluations (procedures, steering experience, quality of reports, impact, lessons learnt, etc.) or carrying internal evaluations (where applicable).
- *Evaluations to assess the quality of monitoring systems and data* are recommended. There are already initiatives undertaken in some countries to evaluate the functioning of monitoring systems with the aim of improving them. This is a good practice example that could be considered by other Member States.

The supply side

1. **External supply of evaluation services.** In some Member States evaluation units carry out internal evaluations, but in most cases the supply of evaluation services remains external to the administrations. In many countries, local private companies tend to dominate the provision of services, with limited involvement of research institutions or foreign expertise. Internal evaluations could be useful for some types of evaluations, in particular in the context of on-going evaluations, but limited capacity still exists in this regard, especially in the EU12.
2. **Differences in evaluation market development.** Research evidence supports the widely-held opinion that evaluation markets in many of the EU12 countries are still under-developed.
3. **Still room for improvement in quality of evaluation services.** While the general opinion - as expressed by both evaluators and administrations - on the quality of external evaluation services is quite positive, there is certainly room for improvement for instance, in areas such as quantitative and experimental techniques, horizontal issues, etc. Training and education opportunities for evaluators are still limited in the EU12.

Recommendations for strengthening the supply side

For Member States

It is more difficult to come up with recommendations to strengthen the supply side since it develops outside the administrative system. However, the following means to reinforce evaluation capacity development on the supply side could be considered.

- The commissioning of *more evaluations* would provide more possibilities for market participants to take part in evaluation activities and thereby contribute to strengthening capacity. There needs to be sufficient demand for evaluation if market participants are to specialise in the provision of such services and invest in strengthening of relevant capacities. This is particularly important in

the EU12 which have a short track record in evaluation. More *smaller scale thematic* evaluations have the potential to raise interest in providing evaluation services and to attract new players into the market.

- National authorities can create incentives for the development of external evaluation expertise by using *tendering policy and requirements* as an instrument (e.g. attempting to widen evaluation community, attract academia or consortia with foreign expertise).
- Member States are encouraged to *publicise evaluation plans* to inform the wider evaluation community as to the amount of evaluation work foreseen.
- Member States are encouraged to consider undertaking so-called *meta-evaluations* of evaluation reports to assess their quality, methods used, results received, utilisation of recommendations and the like. Such an exercise would constitute a learning experience both for administrations and evaluators (if publicised). It would potentially contribute to better quality control and quality of evaluation reports.
- Advertise *evaluation events* organised by public authorities and *invite evaluators* to participate in them.

For evaluation market participants

- Here it could be emphasised that evaluation societies contribute to developing evaluation capacity by providing opportunity for debate and encouraging learning. Thus actions taken to establish evaluation societies in a number of Member States constitute good practice examples.

Dissemination and utilisation

1. ***Dissemination quite good.*** Evaluation reports are quite widely available to the public, including via internet. Accessibility to reports (or at least awareness of them) is higher in the EU15.
2. ***Utilisation of results can be improved.*** Member States have different arrangements in place to address evaluation findings with various functions exercised by monitoring committees, evaluation steering groups, evaluation units or managing authorities. However, the assessment of the influence of the most recent ex ante evaluation, which is closely integrated into the programming process, shows that utilisation of evaluation results is still limited. Research evidence suggests that impact of ex ante evaluations was greatest on the quantification of objectives and indicator selection, with influence on financial allocations being the most limited. Research results also suggest that there exists a credibility gap in the EU12 where administrators are more critical of quality of external evaluation services and where evaluators feel that their results are not sufficiently used and disseminated.

Recommendations for strengthening dissemination and utilisation

For Member States

Although the general availability of evaluation reports is quite good across the EU, there is still room for development in this area, especially in the EU12.

- Such measures as *centralised online access* to Structural Fund evaluation reports can contribute to better awareness of evaluation projects and their higher impact.
- Some Member States have set *minimum requirements for publication* of the evaluation reports (e.g., requirement that at a minimum, the executive summary of evaluation reports shall be made publicly available). This is a practice which could be considered by other countries.

- Where these do not exist, it is recommended to consider setting up *formal procedures to address and follow-up of evaluation results* in order to ensure their better integration into decision-making.

Wider governance environment

1. ***National evaluation practice in most cases limited.*** Administrators and evaluation market participants regard the role of evaluation in the national budgetary and policy-making processes as being of limited importance and note that in most cases evaluation reports are only sometimes used in wider public debates. The assessment of the influence of EU Structural and Cohesion Fund requirements on national evaluation practice and capacity generally ranges from “moderate” to “strong”. Administrators from the EU15 tend to think that the overall administrative and political culture in their country is neutral or broadly favourable to evaluation, while their counterparts from the EU12 are more critical and see administrative and political culture as neutral or somewhat unfavourable to evaluation in public policy making. The opinion of evaluation market participants is similar.

Recommendations in relation to wider governance environment

- For European Commission*** ▪ The European Commission should encourage Member States to *use Structural funding for evaluation capacity building* within wider administration via Administrative Capacity OPs (funded by the ESF in a number of Member States) or through establishing specific objectives in the ESF programmes, which focus on the development of institutional capacity.
- For Member States*** ▪ *Administrative Capacity OPs* (funded by the ESF in a number of Member States) or *ESF programmes* could be used as a vehicle *to enhance evaluation capacity* more widely and make the wider administrative culture (outside the scope of Structural Fund system) more conducive to evaluation practice.

1. INTRODUCTION

1.1. Background

Articles 158-162 of the Treaty establishing the European Communities lay down that the Union should promote an overall harmonious development and strengthen economic and social cohesion by reducing development disparities between the regions. The financial instruments to pursue these objectives are the Structural and Cohesion Funds, currently accounting for one third of the Community budget. The effectiveness of assistance from the Funds is promoted through *inter alia* the incorporation of reliable evaluation into programme design and implementation. The general aim of evaluation of cohesion policy is to "improve the quality, effectiveness and consistency of the assistance from the Funds and the strategy and implementation of operational programmes with respect to the specific structural problems affecting the EU Member States and regions concerned, while taking account of the objective of sustainable development".¹ The Council Regulations require the carrying out of Structural and Cohesion Fund evaluations before, during and after the programming period.²

The new Regulations for the 2007-2013 programming period contain certain obligatory requirements for the evaluation of Structural and Cohesion Funds.³ However the general approach to evaluation has been significantly modified compared to 2000-2006 period: there is a shift away from a concept of mid-term evaluation driven by regulatory imperatives towards an on-going evaluation⁴, which allows for more flexibility as regards evaluation timing, scope and design.⁵ This provides EU Member States with more leeway in conducting Structural and Cohesion Fund evaluations and is intended to encourage evaluations that are adapted to internal demands and that meet the needs of decision makers. At the same time, the relaxation of top-down requirements gives rise to considerable challenges in promoting the overall approach to on-going evaluation and ensuring that the Member States adopt a pro-active stance in undertaking evaluations and use them as an integrated management tool. Adequate administrative capacity, including evaluation capacity, is needed to meet these challenges and to fulfil the promises of the new approach to Structural and Cohesion Fund evaluation.

Against this background, the key question and important driver of this research is to what extent such evaluation capacity exists in the EU Member States and how countries are prepared for new challenges arising from the modified approach to Structural and Cohesion Fund evaluation. It is generally recognised that, even though evaluation capacity in the EU has been consistently

¹ European Commission. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf

² Council Regulation (EC) No 1260/1999 of 21 June 1999 laying down general provisions on the Structural Funds; Council Regulation (EC) No 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999, etc.

³ Member States shall carry out evaluations (1) linked to the monitoring of operational programmes in particular where that monitoring reveals a significant departure from the goals initially set or (2) where proposals are made for the revision of operational programmes. (Article 48(3) of the Regulation 1083/2006)

⁴ On-going evaluation is a process taking the form of a series of evaluation exercises. Its main purpose is to follow on a continuous basis the implementation and delivery of a programme and changes in its external environment, in order to better understand and analyse outputs and results achieved and progress towards longer-term impacts, as well as to recommend, if necessary, remedial actions.

⁵ European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf

developing and that Structural and Cohesion Fund evaluation requirements have played a positive role in driving this process, the level of evaluation capacity across the EU Member States is still uneven.⁶ The EU12 have only recently started to establish their evaluation structures as well as to apply evaluation tools in planning and management of their Structural and Cohesion Fund programmes. But differences exist not only between the EU15 and EU12 Member States, which are essentially a function of duration of EU membership, but exist across the EU27 reflecting the peculiarity of national contexts: “architecture“ of governance within the country (e.g. centralisation / decentralisation), administrative traditions, organisational and cultural features.⁷ This diversity mirrors the complexity of the environment in which Structural and Cohesion Fund programmes are implemented and confirms that evaluation capacity development is a multi-dimensional process, its path of evolution depending very much on institutional traditions in each country as well as internal demands in terms of planning and implementing policy actions.⁸

1.2. Objectives of the Study

Bearing in mind this complex Structural and Cohesion Fund evaluation environment, this study attempts to develop an evaluation capacity benchmarking tool which allows for a systematic assessment and comparison of evaluation capacity development in the EU27 while taking account of different time perspectives and contexts in the individual EU Member States. The development of such a tool is intended to assist those institutions - at the regional, national and EU levels - which undertake efforts to develop evaluation capacity to ensure that evaluation is an effective tool that delivers added value to management of the Structural and Cohesion Fund assistance.

The research has a threefold *aim*, namely:

- (1) *conceptualise evaluation capacity* and establish a framework of *criteria and indicators* to analyse its development within the Structural and Cohesion Fund administrations in the EU Member States;
- (2) *test* the proposed benchmarking framework by carrying out *analysis* of evaluation capacity development within the Structural and Cohesion Fund administrations;
- (3) present *recommendations* on how the opportunities for further organisational learning within the EU Member States administrations responsible for evaluation of Funds can be created.⁹

In order to develop a robust and relevant benchmarking tool, a rounded understanding of evaluation capacity is necessary. Thus the research starts with the conceptualisation and definition of evaluation capacity. On this basis the benchmarking criteria and indicators are developed to analyse evaluation capacity development within the Structural and Cohesion Fund administrations. The proposed benchmarking framework is then applied in an analysis of evaluation capacity development in the EU Member States. This exercise is intended to test the framework, generate data on evaluation capacity within the Structural and Cohesion Fund administrations and provide evidence for refining the benchmarking framework and developing recommendations.

⁶ Mairate, A. „Developing evaluation capacities in the Member States: the case of Structural Funds”, IDEAS Workshop, Prague 19/20 June 2006. <http://www.ideas-int.org/Documents/Mairate%20paper.doc>

⁷ Developing capacity for socio-economic evaluations. - EVALSED on-line resource - <http://www.evalsed.info/page.aspx?id=gui101>

⁸ Mairate, A. „Developing evaluation capacities in the Member States: the case of Structural Funds”, IDEAS Workshop, Prague 19/20 June 2006. <http://www.ideas-int.org/Documents/Mairate%20paper.doc>

⁹ It is important to note that whereas this research project covers evaluation capacity development within Structural and Cohesion Fund administrations in the EU Member States, it does not concern project evaluation (appraisal of projects) – it is focused exclusively on programme evaluation.

The research resulted in a tested and ready-to-use analytical framework that systematically addresses multiple aspects of evaluation capacity development across the EU and allows for a review of progress in this field. The study should be considered as a source of “learning” and guidance and serving as a reference point for discussions on evaluation capacity development at the regional, national and EU level.

1.3. Structure of the Report

The following chapter 2 presents the methodology for conceptualising evaluation capacity and elaborating criteria and indicators to analyse its development. This is accompanied by an outline of the methodology for the application and testing of the proposed framework. Chapter 3 overviews the literature on evaluation capacity development. This leads (chapter 4) to the preliminary version of framework of criteria and indicators to analyse Structural and Cohesion Fund evaluation capacity development in the EU Member States administrations. The results of the framework testing exercise are discussed in chapter 5. The final chapter 6 presents overall conclusions and recommendations.

2. RESEARCH METHODOLOGY

The purpose of this chapter is to provide an overview of the methodological approach utilised in this study, covering the conceptualisation of evaluation capacity and the subsequent development, application, testing and revisions of the benchmarking framework.

2.1. Conceptualising Evaluation Capacity and Developing Benchmarking Framework

The key component of this project is the development of a benchmarking framework which facilitates analysis and comparison of Structural and Cohesion Fund evaluation capacity in the EU Member States. In order to develop a robust and relevant benchmarking tool, a rounded conception of evaluation capacity is necessary. Thus the research started with the analysis and conceptualisation¹⁰ of evaluation capacity. The methodology used in the conceptualisation of evaluation capacity and the development of the benchmarking framework comprised a number of key stages.

Firstly, a comprehensive review of literature in the evaluation capacity area was undertaken. In reviewing the literature, the focus was on the following *issues*:

- *What is evaluation capacity development?*
- What are the key *dimensions* of evaluation capacity development? Put another way, what are the key issues that governments and administration must address if evaluation capacity is to be developed and sustained?
- Are there *criteria and indicators* that can be used to measure the state of evaluation capacity in a particular country under these dimensions?
- In more general terms, what are the *key lessons* (or critical success factors) that emerge from the literature as being important for the successful development of evaluation capacity?

While the development of the framework had to be informed by the latest developments in the evaluation capacity development field internationally, there was also a need to design a framework that was applicable and relevant in the specific field of evaluation of EU Structural and Cohesion Funds. Therefore account was taken of:

- the specific *institutional context* in which the evaluation of EU Structural and Cohesion Funds takes place; and
- the evaluation *challenges* facing the EU administrations over the forthcoming *2007 to 2013* period.

Drawing on the outcomes of the literature review and taking account of the specificities of the evaluation of Structural and Cohesion Funds, a draft evaluation capacity benchmarking framework consisting of criteria and quantitative and qualitative indicators was developed.

2.2. Application of the Benchmarking Framework

Once evaluation capacity was conceptualised and an evaluation capacity benchmarking framework proposed, the next step was to apply it in an analysis of evaluation capacity development within

¹⁰ “Conceptualisation” encompassed the elaboration of a definition of evaluation capacity and analysis of its key dimensions.

Structural and Cohesion Fund administrations in the EU Member States. This element of the research was intended to serve several *objectives*:

- Most importantly, to allow for *piloting and verifying* the appropriateness of the devised benchmarking framework.
- An additional goal was to get a *comparative view of the evaluation situation* in EU Member States, taking due account of national context in the individual Member States.
- Finally, the analysis was used as a *basis for* the development of evidence-based *recommendations* on the application of the framework and further development of evaluation capacity development.

As time and resource constraints did not allow for in-depth analysis of all EU Member States, the focus was on testing the proposed benchmarking framework and assessing its appropriateness. In other words, the application of the framework was not intended to be a full benchmarking exercise but rather a development exercise to elaborate the framework and test its applicability in different countries and via different methods.

A combination of survey and case study methods were used to collect information on Structural and Cohesion Fund evaluation capacity and to test the applicability of benchmarking framework. Data collection methods are presented in detail below.

2.2.1. Survey

Over the period 30 May to 13 June 2007 a cross-national survey covering all 27 EU Member States was conducted on the basis of self-administered questionnaires. The aim was to reach a large number of stakeholders in all EU Member States in a short period of time, collect comparable primary data and test the evaluation capacity indicators (i.e. to see whether these are manageable through a survey and capable of producing relevant data, facilitating comparison and identification of good practice).

Two key groups of *stakeholders* were surveyed:

- (1) The prime target group were *administrators* with responsibility for Structural and Cohesion Fund evaluation. The survey covered administrators both at *national level* as well as at *programme level*.
- (2) In addition, *external organisations* (companies, universities/research institutions and other non-governmental organisations) working in evaluation and having experience in undertaking Structural and Cohesion Funds evaluations over the period 2000-2007 were surveyed.

Full respondent list comprised 446 contacts for Structural and Cohesion Fund administrations and 509 contacts for evaluation market participants. The table 1 below specifies the number and type of respondents by country.

Table 1. Questionnaire recipients by country and type

EU Member State	Structural and Cohesion Fund administrators	Companies	Academic institutions	Other (NGOs)	Total
Austria	14	9	2		25
Belgium	11	10	7		28
Bulgaria	12	3	1		16
Cyprus	5	2			7
Czech Republic	31	34	13	2	80
Denmark	6	2	4		12
Estonia	9	9	5		23

EU Member State	Structural and Cohesion Fund administrators	Companies	Academic institutions	Other (NGOs)	Total
Finland	9	15	13		37
France	38	17	6		61
Germany	22	19	16		57
Greece	37	32			69
Hungary	11	17	4		32
Ireland	7	7	4		18
Italy	29	17	3	1	50
Latvia	15	9	3		27
Lithuania	15	10	4		29
Luxembourg	5	4			9
Malta	4	5			9
Netherlands	16	8	3		27
Poland	16	15	18	4	53
Portugal	25	22	21		68
Romania	15	11	3		29
Slovakia	19	4	2		25
Slovenia	6	3	4	1	14
Spain	23	12	2	1	38
Sweden	17	12	5	1	35
United Kingdom	29	40	8		77
Total	446	348	151	10	955

As the table 1 indicates, the number of contacts for individual countries varied. This disproportion can be largely explained by variance in the number of staff responsible for Structural and Cohesion Fund evaluation across the EU Member States (and the complexity of Structural and Cohesion Fund programming structures) and in the number of evaluations carried out (and consequently the number of contacts for the supply side). However, it is worth reiterating that the emphasis of this project was not on undertaking a full benchmarking exercise as such but on developing and testing the benchmarking framework. Therefore the representativeness of the sample was not a first-order issue in this case. Rather, the intention was to address the relevant questions to different stakeholders in each EU Member State to find out to what extent different indicator data is available and possible to collect.

Taking into account the fact that the survey had to address several categories of stakeholders, two questionnaires were developed – one addressed to administrators and the other to evaluation market participants. Survey questionnaires are enclosed in Annex 1. Both factual and opinion questions were included in the questionnaires reflecting the different types of indicators (quantitative and qualitative). The questionnaires included both closed and open-ended questions. The balance was towards the closed questions, in order to facilitate analysis and comparison. Another consideration was to facilitate respondents in completing the surveys in as short a time as possible thereby encouraging a better return rate. Several types of responses to the questions were envisaged depending on the type of question. Response options to close-ended questions included dichotomies (yes/no) and multiple choice with one or several possible responses. For opinion questions, rating scales were used. Open-ended questions mainly related to “hard” data.

2.2.2. Country Case Studies

The survey was complemented by four case studies of selected countries. Case studies were intended to serve the following *purposes*:

- The collection of detailed information and *evidence against* various *evaluation capacity indicators*, especially those indicators which could not be covered by survey;

- Facilitate *verification* of the benchmarking framework;
- *Contextualise* data generated through survey and enable the analysis of “how” and “why” questions, especially in relation to good practice cases.

Four EU Member States – Hungary, Lithuania, Romania and England - were selected for an in-depth analysis of Structural and Cohesion Fund evaluation capacity. This selection was guided by several considerations. Given the purpose of the assignment, countries regarded as advancing in evaluation capacity were of particular interest. Good practice examples are interesting as they can help pinpoint success factors and offer potential for learning. In order to test the applicability of benchmarking framework in different settings, the case study countries were selected with regard to the duration of EU membership and size. For similar purposes countries eligible under the Objective 1/Convergence and Objective 2/Regional Competitiveness and Employment objectives were covered. As significant level of Structural and Cohesion Funds is allocated to the EU12, the balance was towards states where aid plays a relatively larger role. Consideration was also given to avoiding the overlap with the case studies that were being carried out for the purposes of updating the EVALSED, the online and interactive resource for the evaluation of socio-economic development (referred to subsequently as the EVALSED).¹¹

Fieldwork in the selected countries was carried out during May - July 2007. Case studies drew on several data collection methods and on accounts of different stakeholders. Each case study involved the selection and review of relevant documentation (desk research), analysis of survey results, fieldwork in the selected countries and telephone interviews. (Full case study reports of individual countries are presented in Annex 2 to this Report.)

Desk research

Desk research was used to collect and analyse available factual information related to Structural and Cohesion Fund evaluation issues, gain background knowledge, determine what additional information needed to be collected from key informants and prepare for the fieldwork missions. A bibliography is provided in each case study report.

Interviews

Primary data was collected during 1 to 2 days long fieldwork missions in the selected countries. Fieldwork missions involved carrying out face-to-face semi-structured interviews with key informants. Interviews complemented and facilitated interpretation of information collected through the survey and desk research and allowed for the exploration of the views of the actors involved.

Interview questions were developed on the basis of evaluation capacity benchmarking framework and linked to its component criteria and indicators. Different types of *stakeholders* were approached and *interviewed*:

- *Individuals working at evaluation units and/or managing authorities at national level* and individuals responsible for ESF evaluation. In some instances stakeholders at programme level have been interviewed.

The primary focus of interviews with the Structural and Cohesion Fund administrators was on the overall architecture of the Structural and Cohesion Fund evaluation system and on the utilisation of evaluation results. Interviews explored in detail such issues as the organisation, management and coordination of the evaluation function; the distribution of responsibilities among institutions, legal basis and operation in practice; advantages and shortcomings of the approach taken; lessons learnt and possible challenges in 2007-2013 programming period. Another line of questioning concerned the linkages between evaluation and other functions such as programming and monitoring and the

¹¹ For this reason Ireland, Czech Republic, Italy, Netherlands, Poland and Spain were not considered as case studies for this research.

utilisation of evaluation results. A significant issue that was also covered during the interviews was evaluation resources, including staff competence and approach to evaluation capacity development in the public administration. In addition, views on the adequacy of evaluation expertise were sought. Finally, aspects of the wider administrative and political environment were discussed in order to establish to what degree it encouraged or hindered the development of Structural and Cohesion Fund evaluation capacity.

- *Private sector evaluators and academics* with experience in evaluation of Structural and Cohesion Fund programmes.

Interviews with the external evaluators and academics focused on such issues as experience and knowledge in Structural Fund evaluations and the situation as regards competition in the evaluation market. Data sources used in evaluation were discussed, including both the data generated by monitoring systems and the quality and availability of up-to-date socio-economic statistical data. An important line of questioning concerned the supply of training or education in the field of evaluation and frequency of evaluation events. Where applicable, involvement in the professional organisation of evaluators was discussed. The views of evaluators on the efficiency of evaluation management and on issues such as the clarity of evaluation terms of reference, were sought.

2.3. Revision of the Benchmarking Framework

On the basis of fieldwork results, strengths and weaknesses of the evaluation capacity benchmarking framework could be identified and necessary revisions of criteria and indicators could be detected. The fieldwork was expected to reveal which indicators were ambiguous, excessive, too complex or of limited relevance for the analysis, comparison and identification of good practices in evaluation capacity development. Data collection feasibility was also taken into account. As a result of the testing exercise, improvements to the framework could be made by omitting excessive benchmarking criteria/indicators and those which proved to be of limited relevance. At the same time ambiguous or complex indicators could be further specified and new ones identified.

2.4. Developing Recommendations

The fieldwork formed a basis for the revision of the framework and recommendations for its future application and use. It also provided evidence against which recommendations on further opportunities for the evaluation capacity development within Structural and Cohesion Fund administrations could be proposed. Recommendations address weaknesses and risks identified in evaluation capacity development and are targeted at different stakeholders.

3. EVALUATION CAPACITY DEVELOPMENT LITERATURE REVIEW

An initial point of departure in conceptualising evaluation capacity and identifying the main elements of the framework to analyse its development was to consider existing literature on this subject. This chapter summarises a review of literature in the field of evaluation capacity development, sometimes referred to as evaluation capacity building or simply as evaluation capacity. The review is organised around a number of themes as follows:

- Section 3.1. considers the *definitions of evaluation capacity development* advanced in the literature as well as issues around what might be termed the *rationale* for evaluation capacity development;
- Section 3.2. summarises a number of important *dimensions of evaluation capacity* identified in the literature;
- Section 3.3. focuses on *evaluation capacity criteria and indicators* against which evaluation capacity can be assessed, outlining a number of classifications that have been developed in the literature; and
- Section 3.4. concludes by summarising the key lessons or *critical success factors* emphasised in the literature as being important in the successful development of evaluation capacity.

3.1. Definition of Evaluation Capacity and its Rationale

3.1.1. Definition of Evaluation Capacity

At the outset, it is important to differentiate the practice of evaluation from evaluation capacity development. Evaluation (defined, for example, as “the systematic assessment of the operation and/or the outcomes of a programme or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy”)¹² is essentially a tool or an instrument, whereas evaluation capacity development constitutes the *process* of setting-up the necessary systems and infrastructures to undertake evaluation. As the World Bank Operations Evaluation Department (2002) puts it: “<...> Monitoring and evaluation are tools, and evaluation capacity development is the process of setting-up a country-based system to conduct and use monitoring and evaluation”.

A number of definitions of evaluation capacity, varying in terms of breadth and complexity, are evident from the literature review. For example, Mackay (1999) defines evaluation capacity development in a rather broad, but nevertheless straightforward, manner as “the development of national or sectoral evaluation systems”.¹³ The definition used by Picciotto (1998) is somewhat broader - “the ability of public institutions to manage information, assess programme performance, and respond flexibly to new demands” - in that it does not confine itself to evaluation as such. Boyle and Lemaire (1999) link their definition of evaluation capacity development to the notion of an evaluation regime, which they describe as “the configuration of evaluation capacity, evaluation practice, organisational arrangements and institutionalisation”. Evaluation capacity development is then defined as “activities and initiatives taken to implement the regime”. The definition presented in the EVALSED is comprehensive in scope: they define evaluation capacity development as “the

¹² Weiss, Carol H., *Evaluation* (1998), New Jersey: Prentice-Hall

¹³ Mackay is the Evaluation Capacity Development Coordinator at the World Bank.

institutional, human, resource, skill and procedural base for conducting evaluations in public policy and public management systems”.¹⁴

Stockdill, Baizerman and Compton (2002) devote a chapter to the issue of conceptualising and defining evaluation capacity building. Initially, they present a conceptual definition of evaluation capacity building: “a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites”. The authors emphasise that evaluation capacity building, contrary to the actual practice of evaluation, is a continuous process. In a colourful metaphor, they characterise evaluation capacity building as “never ending – like the circus performer who constantly spins the many plates on top of the pole she holds”. This discussion leads them to what they term a working definition of evaluation capacity building as “the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine”.

In summary, evaluation capacity development is concerned with creating and sustaining the factors that support evaluation in the public or government sector. In the next chapter, we draw on the various conceptualisations of evaluation capacity development summarised here in arriving at a working definition of evaluation capacity development for the purposes of this project.

3.1.2. Rationale: Why Evaluation Capacity Development Matters

Having discussed definitions of evaluation capacity development, this section reviews briefly what the literature has to say on the rationale for evaluation capacity development. Why is evaluation capacity development important? MacKay (1999) highlights the role of evaluation capacity development as an important aid to sound governance and as a way of helping to achieve high levels of public sector performance. He argues that evaluation capacity development ensures that evaluation findings are available to *assist countries in four key areas* as follows:

- As an input to *resource allocation and budgetary decisions* by government;
- As a *management tool* which reveals the performance of ongoing activities at various levels in turn leading to learning and improvement;
- As an *accountability mechanism*, i.e., holding managers and governments to account for their performance; and
- Demonstrating the *extent* to which development *activities* have been *successful*.

Boyle and Lemaire (1999) pose the question: Why build evaluation capacity? Their “basic response” to this question “is that governments build national evaluation systems because they believe such a system will help them improve their means and methods of governance”. They view evaluation and evaluation capacity development as central elements of a “results-oriented public sector management” (or new public management) approach adopted in many countries over the preceding ten to fifteen years. However, they stress that evaluation is but one component of this broad managerial framework.

More broadly, the EVALSED Guide sees the development of evaluation capacity as part of a wider need to build institutional and administrative capacity (increasingly itself an explicit objective of socio-economic development policy). This need or goal stems from the twin “imperatives” of, firstly, overcoming inefficiencies in traditional public administration (by moving towards a new public management approach) and, secondly, opening up the policy making process to civil society and citizens generally in the pursuit of democratic governance. The authors argue that the diffusion

¹⁴ EVALSED: the on-line resource for the evaluation of socio-economic development <http://www.evalsed.info/glossaryDefinition.aspx?id=66>

of evaluation can contribute to both these imperatives. On the former, they highlight the role of evaluation as a “powerful tool for modernisation in the public sector, for cost reduction and for greater responsiveness to citizens”. On the latter, the practice of evaluation, which opens up programmes to the scrutiny of external stakeholders and citizens, is regarded as in itself “an embodiment of the principles of democratic governance”.

In summary, evaluation capacity development is important from two main perspectives:

- As a contributor to better public sector performance and enhanced administrative capacity; and
- As a way of improving governance generally.

Having acknowledged the importance of the evaluation capacity development, the following section explores main dimensions of evaluation capacity.

3.2. Key Dimensions of Evaluation Capacity

3.2.1. Introduction: The Basic Evaluation Capacity Development Framework

In this section, the focus is on the key dimensions of evaluation capacity as identified in the evaluation literature. Another way of looking at this is to ask what are the key, strategic issues that governments and administrations (such as, for example, Structural and Cohesion Fund administrations in the EU Member States) must address if evaluation capacity is to be successfully built?

A number of different frameworks, made-up of these critical evaluation capacity development dimensions, can be observed in the literature. There is a strong degree of consensus in the literature around these key dimensions although this is sometimes obscured by the differences in the terminology employed by experts in the field. For example, terms such as „key issues/decisions“, „dimensions“ or „building-blocks“ generally refer to the main components of evaluation capacity development. (We will use the term „dimensions“.) Below this level, the terms „criteria“ and „indicators“ are often used.

A good place to start is with the analysis of the development of evaluation capacity contained in the EVALSED. The authors identify the following as what they term “the most important *key decisions* to be made when starting *to develop evaluation capacity* in a strategic way”:

- *Architecture*: this is concerned with the location, structuring and coordination of evaluation functions;
- *Strengthening evaluation demand*: this is about “ensuring that there is an effective and well managed demand for evaluations”;
- *Strengthening evaluation supply*: concerned with issues around evaluation skills and competencies with appropriate organisational support;
- *Institutionalising evaluations*: the issue in question is “the building in of evaluation into policy making systems and across the broader policy system”.

Most of the above dimensions – architecture, demand, supply and institutionalisation – are common to other evaluation capacity development frameworks found in the literature. For example, many of the country-specific analyses of evaluation capacity development by the World Bank focus explicitly on three of these dimensions, namely demand, supply and architecture.¹⁵ The fourth dimension – institutionalisation – is also referred to, but more as an overarching aspect that captures the outcome of the three other dimensions (in other words, the interaction of demand,

¹⁵ See for example the World Bank’s Operations Evaluation Department ECD Working Paper No. 5 (*Evaluation Capacity Development Comparative Insights from Columbia, China and Indonesia*) published in January 1999.

supply and architecture determines the extent to which evaluation is institutionalised). The World Bank's framework is also drawn on in the paper by Jaszczolt *et al.* (2003) which focuses on demand, supply and institutionalisation framework factors.

In a similar vein, drawing on the work of the International Evaluation Working Group on Policy and Programme Evaluation (which studied evaluation capacity development over a number of years), Boyle, Lemaire and Rist (1999) identify *seven key issues* which they say “must be dealt with by governments wishing to enhance the demand for and supply of evaluation”, and thus enhance evaluation capacity. These issues and the key questions relating to each are set out in table 2 below.

Table 2. Key Issues and Questions in Building Evaluation Capacity

Issue	Question
Anchoring the evaluation regime	Should the evaluation regime be anchored in the legislative or the executive branch, or both, and where within those branches?
Anchoring evaluation capacity within organisations	What are the advantages and disadvantages of: (a) centralised versus decentralised evaluation capacity (b) using internal versus external evaluators?
Evaluation coverage	What kind of activities should be covered by evaluation?
Linking evaluation with other public sector functions and institutions	How and to what extent should evaluation interact with budgeting, auditing, strategic planning and monitoring?
Using evaluation in decision-making	What are the key domains of use, and how can evaluation utilisation be improved?
Professionalizing the evaluation function	To what extent should evaluation be a discipline in its own right, and what skills and competencies are needed by evaluators?
Fostering demand	What are the most effective ways of fostering demand for evaluation activity?

Source: Boyle, Lemaire and Rist (1999)

The above issues and associated questions can be understood in terms of the dimensions highlighted in the EVALSED and World Bank frameworks considered earlier. For example, the first two issues concerning the “anchoring” of evaluation capacity constitute “architectural” issues. The issues of “evaluation coverage” and “using evaluation” are questions of institutionalisation. The professionalisation of the evaluation function is concerned with evaluation supply while evaluation demand („fostering demand“) also features in the Boyle et al. framework.

One other framework worthy of note is that presented recently by Lahey (2007).¹⁶ His framework rests on four “*essential building blocks*” as follows:

- *Vision*; this encompasses an understanding of the role of monitoring and evaluation information in public sector management and decision-making and strategic leadership;
- *Enabling Environment*; this captures the commitment to sustain a monitoring and evaluation system over the long-term and the political will to support underlying values and ethics such as objectivity, transparency and good governance;
- *Infrastructure to Supply Monitoring and Evaluation Information*; in other words the supply-side dimension;
- *Infrastructure to Demand and Use Monitoring and Evaluation Information*; the demand-side dimension.

Again, Lahey's “building blocks” correspond closely to the key evaluation capacity development dimensions or issues emphasised in the other contributions outlined above. Lahey's “Vision”

¹⁶ *A Framework for Developing a Monitoring & Evaluation System*, Presentation by Robert Lahey (REL Solutions Inc.) at the International Conference on Monitoring and Evaluation, Kiev, Ukraine, February 2007.

building block captures elements relating to leadership that in other frameworks are associated with demand or institutionalisation dimensions.

In sum, a strong degree of common emphasis as regards the key dimensions of evaluation capacity development can be observed from the literature. In the following sub-sections, each of the four main dimensions (as presented in the EVALSED) are considered in more detail.

3.2.2. Evaluation Demand

The demand side refers to the commissioning side of the evaluation process. It is concerned with factors to do with the capacity and commitment of governments to commission and use evaluation. Most commentators agree that a demand for evaluation is an essential starting-point in building evaluation capacity development. Unless governments are persuaded as to the benefits of evaluation, neither pressure from external donors nor the availability of a supply-side capacity will result in the institutionalisation of evaluation in the business of government (Mackay, 2006).

Mackay (2006) notes that a barrier to evaluation demand often stems from a lack of knowledge of what evaluation entails. He refers to the frequent occurrence of a “chicken and egg problem whereby a lack of government demand for evaluation stems from a lack of understanding of evaluation and what it can provide, this lack of understanding stems from a lack of evaluation experience which, in turn, is itself explained by weak demand”. He argues that the way around this dilemma is to try to increase awareness of evaluation and its uses.

In terms of how effective evaluation demand can be fostered, the literature generally differentiates between approaches that rely on incentives (or „carrots“ as Toulemonde refers to them in Boyle and Lemaire, 1999) or alternatively - on decree or regulation-based approaches (what Toulemonde refers to as „sticks“). For example, obligations deriving from EU regulations as well as the impact of relevant Commission guidelines play an important role in shaping evaluation practice in the EU Structural and Cohesion Funds context.

3.2.3. Evaluation Supply

The supply side focuses on whether the necessary resources, skills and information infrastructures are available to effectively respond to demand pressures. Guerrero (1999) argues that evaluation supply requirements are a function of the questions that evaluations must answer. He identifies four interdependent categories of supply requirements or resources: methods and standards, information, financial resources and professional skills.

A key point stressed in the literature is the interdependence of demand and supply. Picciotto (1995) stresses that demand and supply must be nurtured in parallel on the basis that “Say’s law does not apply fully: good evaluation does not always create its own demand”.¹⁷ Feinstein (2002) argues that the inverse (Yas’ law – that demand induces supply) only applies to a limited extent. Thus most commentators stress the need to work in parallel on both the demand and supply sides in the development of evaluation capacity.

3.2.4. Evaluation Architecture

The architecture dimension refers to how evaluation systems are organised and coordinated. There is a strong degree of overlap with demand considerations in that the way in which evaluation

¹⁷ In economics, Say’s Law is a principle attributed to French businessman and economist Jean-Baptiste Say (1767-1832) stating that there can be no demand without supply.

systems are organised will strongly influence the level and nature of evaluation demand. As highlighted by Boyle, Lemaire and Rist (1999), important issues here concern the branch of government that the evaluation function should be located in and the advantages/disadvantages of centralised versus decentralised and internal versus external evaluation approaches (see Table 1 above). Another key issue, also highlighted by Guerrero (1999), is the linkages between evaluation and other related public sector functions such as budgeting and audit. The EVALSED notes that in the early stages of the development of evaluation capacity, ownership of an evaluation capacity development strategy is generally located in a dedicated unit within central government such as in a Treasury Ministry, Prime Ministers office or a department responsible for the EU Structural Funds. A point made in EVALSED is that if evaluation is to become a valued function within the public sector “it must itself mirror the architecture of governance within the country concerned”.

3.2.5. Institutionalisation

As noted above, the institutionalisation dimension can be viewed in two ways. In its narrower manifestation, it can be thought of as a dimension of evaluation capacity in its own right with a focus on evaluation utilisation and follow-up mechanisms. More broadly, institutionalisation can be seen as an overarching factor, which reflects the interaction of the other evaluation capacity development dimensions, and captures the extent to which evaluation forms an integral part of the policymaking process and of the business of government generally. As Mackay (2006) puts it “<...> by institutionalisation I mean the creation of a monitoring and evaluation system which produces monitoring information and evaluation findings which are judged valuable by key stakeholders, which are used in pursuit of good governance and where there is sufficient demand for the monitoring and evaluation function to ensure its funding and its sustainability for the foreseeable future”.

Thus, a focus on institutionalisation allows for consideration of the wider, cultural factors that determine the degree of influence of evaluation. An important contribution to this debate stems from the work of Weiss (1999), who hypothesises on the national conditions that promote application of evaluation to the business of government. In the first instance, she highlights two general conditions - the existence of a democratic and competitive political system and the decentralisation of policy - as capturing a single construct “that is, the openness of the system to evidence and argument”. She argues that this “system permeability” is an important feature of those *regimes* that make use of evaluation. Other important *features* include:

- The *functional specialisation of policymakers*; i.e., when policy makers are specialised experts in their policy domain they will have been exposed to the “understandings and insights provided by evaluation”;
- The *educational and professional background of policy makers*; in particular she argues that wherever law is the dominant profession of legislators and top civil servants, the “social sciences seem to fare relatively poorly”;
- On the supply-side (or the “sending” side as she refers to it), the presence of a *thriving social science community* or communities and, within this, a university system that is hospitable to the social sciences;
- Again, on the supply-side, the availability of a *sizeable group of social scientists* who are interested in conducting policy-oriented research;
- The presence of *institutions that bridge the academia-government gap*;
- A “*climate of rationality*” (see discussion above); and

- The *nature of the issues on the policy agenda*, the characteristics of which will determine the extent to which they lend themselves to evaluation influence. In particular, the issues need to be ones where evaluators have worked and not issues that are emotive or highly-charged.

3.2.6. Summary and Conclusion

The evaluation capacity development “framework” presented in the EVALSED, comprises four dimensions relating to evaluation demand, supply, architecture and institutionalisation. Most, if not all, of these dimensions are also highlighted by other evaluation capacity development experts and organisations, including the World Bank. Thus a reasonable conclusion to be drawn is that there is a strong consensus in the literature as to the strategic issues that must be addressed by governments if they wish to build evaluation capacity. In the next chapter, we integrate most of these dimensions into our Structural and Cohesion Funds capacity framework.

A second important point of note is the strong degree of overlap and interdependence between the various evaluation capacity development dimensions. The interdependence of evaluation demand and supply and the importance of developing both in parallel has been noted. There is also a strong degree of overlap between the organisation of the evaluation function (architecture) and the nature of evaluation demand.

A final point is that it is important to take account of the wider factors that influence the role of evaluation in policymaking generally. Some of these wider, cultural-type factors have been highlighted under the institutionalisation dimension above. These are taken into account when developing the evaluation capacity benchmarking framework.

3.3. Evaluation Capacity Criteria and Indicators

3.3.1. Introduction

In this section, a number of evaluation capacity development criteria and indicator typologies from the literature are reviewed. These are criteria or indicators that have been developed by various experts for the purpose of the assessment of evaluation capacity mainly at country or government level. This is of obvious relevance to our objective of constructing an evaluation capacity development benchmarking framework for the Structural and Cohesion Fund administrations in the EU Member States. In essence, the criteria and indicators can be understood as breakdowns or more detailed expressions of the evaluation capacity development dimensions discussed in the section above even if not presented as such in the literature.

3.3.2. Furubo and Sandahl Benchmarking Criteria

One set of evaluation capacity criteria stems from the work of Furubo and Sandahl (2002), who benchmarked the maturity of evaluation cultures across 21 countries, including EU10. Their analysis was carried out on the basis of 9 criteria¹⁸ which are set out below in table 3 together with our view as to the relevant evaluation capacity development dimension (in a separate column):

¹⁸ These nine criteria were subsequently refined and expanded into the so-called “9+3” criteria under the aegis of an *Evaluation Advisory Group* (established in 2003) which included experts from the European Commission, EU Member States and the candidate countries. Three additional criteria were added to adapt the framework to the specific needs of EU Candidate countries:

1) There is monitoring capacity (Supply);

Table 3. Furubo and Sandahl Benchmarking Criteria

No.	Evaluation Culture Benchmarking Criteria	Relevant ECD Dimension(s)
1.	Evaluation takes place in many policy domains	Demand
2.	There is a supply of evaluators specialising in different disciplines who have mastered different evaluation methods and who conduct evaluations	Supply
3.	There is a national discourse concerning evaluation in which more general discussions are adjusted to the specific national environment	Institutionalisation
4.	There is a profession with its own societies or frequent attendance at meetings of international societies or at least some discussion concerning the norms or ethics of the profession	Supply/Institutionalisation
5.	Institutional arrangements in the government for conducting evaluations and disseminating their results to decision makers	Architecture/Institutionalisation
6.	Institutional arrangements are present in Parliament for conducting evaluations and disseminating them to decision makers	Architecture/Institutionalisation
7.	An element of pluralism exists, that is, within each policy domain there are different people or agencies commissioning and performing evaluations.	Demand/Supply
8.	Evaluation activities within the Supreme Audit Institution.	Demand
9.	The evaluations done should not just be focused on the relation between inputs/outputs or technical production.	Demand

Source: Furubo, Rist and Sandahl (2002)

Of the nine criteria, three can be designated as relating to demand, one to supply and one to institutionalisation dimensions. The remaining four criteria are each linked to two evaluation capacity development dimensions. For example, criteria 5 and 6 (relating to evaluation arrangements in government and Parliament respectively) capture aspects of both the architecture of the evaluation system and the extent to which evaluation is institutionalised. Criterion number 7 refers to both the commissioning and performing of evaluations and for that reason is linked to both demand and supply.

3.3.3. EVALSED Evaluation Capacity Indicators

An interesting perspective on evaluation capacity developed in the EVALSED Guide is the need to consider the issue of evaluation capacity at different levels. The EVALSED argues that “Evaluation capacity is multi-faceted and needs to be located at many different levels that reinforce each other”.¹⁹ In this framework, evaluation capacity needs to be considered at four levels: individual; organisational; inter-organisational and at societal level. The EVALSED sketches a number of

2) There is diversity of evaluation: strategy, policy, programme and projects (Demand);

3) There is an information flow within government relating to evaluation (Architecture).

¹⁹ <http://www.evaled.info/page.aspx?id=gui=100>

indicators or arrangements that one might expect to find at each of these levels which are presented in the table 4 below:²⁰

Table 4. Evaluation Capacity Indicators at Different Levels

Level	Indicator or Arrangement	Relevant ECD Dimension
Individual	▪ People throughout government who have experience and skill in evaluation which is regularly renewed and reinforced	▪ Demand
	▪ Availability of training options delivered by various providers	▪ Supply
Organisational	▪ Evaluation routinely undertaken at each policy and programming stage	▪ Demand/Institutionalisation
	▪ Evaluation findings integrated into decision-making	▪ Institutionalisation
	▪ Managers perceive evaluation as an important input	▪ Institutionalisation
	▪ Regular flow and coverage of evaluations	▪ Demand
	▪ Follow-up procedures	▪ Architecture/Institutionalisation
	▪ Procedures to accumulate evaluation findings	▪ Architecture/Institutionalisation
Inter-organisational	▪ Coordination mechanisms (via networks or functions)	▪ Architecture
	▪ Evaluation requirements embodied in legislation/policy/regulatory activity	▪ Demand
	▪ Well-defined market	▪ Supply
	▪ Culture of evaluation	▪ Institutionalisation
Societal	▪ Open and systematic dialogue between policy makers and evaluation specialists	▪ Institutionalisation
	▪ Existence of an “evaluation community of practice”	▪ Supply/Institutionalisation
	▪ Existence of evaluation associations	▪ Supply/Institutionalisation
	▪ Evaluations routinely used by stakeholders	▪ Institutionalisation

Source: EVALSED

Not surprisingly, most, if not all of the indicators, can be related back to the four key dimensions of evaluation capacity articulated in the EVALSED. In the third column of the table 4 above, we identify the relevant dimension (or dimensions in a few cases). Thus, as might be expected, there is a strong degree of consistency as between the key evaluation capacity development issues (discussed above at section 3.2.1) and indicators developed in the EVALSED. Of the 16 indicators, 3 are classified as demand, 2 as relating to supply considerations, 1 as architecture and 5 as capturing institutionalisation aspects. 5 of the indicators are classified as relevant to two dimensions with institutionalisation featuring in each case. This analysis again highlights the extent of overlap

²⁰ <http://www.evalsed.info/page.aspx?id=gui=100>

as between the main dimensions of evaluation capacity development and the overarching nature of the institutionalisation issue.

3.3.4. World Bank Performance Indicators for Evaluation Capacity Building

Mackay (2002) presents an evaluation capacity building “results chain”, comprised of activities, outputs and outcomes with an associated set of suggested performance for each stage. The indicators pertaining to the output and outcome stages are reproduced in the table 5 below.

Table 5. Suggested Performance Indicators for Evaluation Capacity Building

Results Area	Suggested Performance Indicators	Relevant ECD dimensions
Outcomes	<ul style="list-style-type: none"> ▪ Monitoring and evaluation (M&E) findings are used in budget decision making, in sectoral strategy making and in line management decisions 	<ul style="list-style-type: none"> ▪ Institutionalisation
	<ul style="list-style-type: none"> ▪ M&E findings are used by media, in parliamentary debates, and in NGO submissions to government 	<ul style="list-style-type: none"> ▪ Institutionalisation
	<ul style="list-style-type: none"> ▪ Government structures and processes have been realigned to commission M&E findings and to feed them into budget processes and into ministries’ planning and management processes 	<ul style="list-style-type: none"> ▪ Institutionalisation/Architecture
Outputs	<ul style="list-style-type: none"> ▪ Greater quantity and better quality of monitoring information and evaluation findings 	<ul style="list-style-type: none"> ▪ Supply
	<ul style="list-style-type: none"> ▪ Formal M&E framework is established by government 	<ul style="list-style-type: none"> ▪ Demand
	<ul style="list-style-type: none"> ▪ Number of officials who undertake M&E training 	<ul style="list-style-type: none"> ▪ Demand
	<ul style="list-style-type: none"> ▪ Number of officials working on M&E (full-time equivalents) 	<ul style="list-style-type: none"> ▪ Demand
	<ul style="list-style-type: none"> ▪ Number of evaluations or reviews conducted 	<ul style="list-style-type: none"> ▪ Demand

Source: MacKay (2002)

The three outcome indicators all essentially relate to the extent to which evaluation is institutionalised with the third indicator (realignment of government structures and processes) also capturing aspects of the architecture or organisation of the evaluation system. The first output indicator is concerned with evaluation supply while the remaining four all relate to demand.

3.3.5. Summary and Conclusion

A strong degree of commonality can be observed across the various sets of evaluation capacity criteria and indicators outlined above. For example, all frameworks capture demand-side factors such as the coverage or extent of evaluation activity. On the supply-side, factors such as the presence of suitably qualified evaluation expertise and the existence of evaluation societies or associations are captured. A noticeable feature is that many of the typologies include a range of sometimes quite broadly-defined indicators that essentially relate to the institutionalisation of evaluation.

In general, the criteria and indicators can be linked back to one or more of the key dimensions of evaluation capacity (demand, supply, architecture, institutionalisation) outlined in the preceding section, even if not presented in these terms by their respective authors. A significant number of the

criteria and indicators appear to relate to more than one of the dimensions which suggests a degree of overlap as between the four dimensions.

From the perspective of the benchmarking focus of this study, the findings are encouraging in that it would appear feasible to develop a framework based on some general core evaluation capacity dimensions with a series of more specific or detailed criteria and indicators. However, because most of the indicator frameworks refer (implicitly if not explicitly) to evaluation capacity at the level of government or society as a whole, they would require adaptation in the specific context of this study which is concerned with evaluation capacity within the Structural and Cohesion Fund administrations in the EU Member States. For example, some of the elements included in the frameworks are of limited relevance to capacity within the Structural and Cohesion Fund administrations – such as the criteria in the Furubo and Sandahl approach, relating to evaluation roles and functions in Parliament and Supreme Audit Institutions (except insofar as they provide insights into the wider administrative environment or culture within which the Structural Funds administration operates, for example, the criterion in the Furubo/Sandahl framework relating to the extent of evaluation across policy domains). Therefore, while none of the discussed frameworks could, in their entirety, be adopted for the purposes of this study, they help in informing the selection of criteria and indicators of evaluation capacity at the level of the Structural and Cohesion Funds administrations. This will be presented in chapter 4. But before that, the main lessons that emerge from the literature as crucial for the successful development of evaluation capacity are considered.

3.4. Critical Success Factors for Evaluation Capacity Development

In this section, the focus is on the main findings (or lessons) from the literature as to the factors that are critical to the building of effective evaluation capacity. This complements and “rounds-off” the more specific review of evaluation capacity development dimensions and indicators in the preceding sections.

A useful place to start is with the four key themes that emerge from the comprehensive analysis of evaluation capacity development in *Boyle and Lemaire (1999)*. In a postscript entitled “Evaluation Capacity Building – A Journey without an End”, they firstly acknowledge that the practice of enshrining evaluation within public sector organisations “presents a significant challenge to governments and bureaucrats alike”. They then highlight the *four key themes* that emerge from the various contributions to their publication:

- A first lesson is the importance of a *balance between the demand for and supply* of evaluation capacity;
- A second is the *need for flexibility in utilising opportunities* and incentives to foster evaluation capacity development;
- A third and perhaps overarching lesson that emerges from their analysis (although not presented in these terms by the authors) is that “In essence, what is needed is an *approach that recognises the complexity of demand for evaluation, promotes coherent demand strategies, and adjusts supply strategies accordingly*”;
- A final lesson concerns the *mainstreaming of evaluation into government functions* with the authors finding that “For demand to be strong and effective requires links within and between the different markets for evaluation and the different public sector functions”.

Other points emphasised by Boyle and Lemaire concern the importance of building evaluation systems that are *resilient* and for evaluation regimes to be *flexible* to meet the various market demands which exist for evaluation.

Mackay (2006) notes that the literature on the strengthening of government monitoring and evaluation systems “suggests a broad agreement of a number of *key lessons*”. He highlights eleven such lessons:

- (1) *Substantive government demand* is a prerequisite for successful institutionalisation of monitoring and evaluation systems.
- (2) *Incentives* are an important part of the demand side and, in particular, there need to be strong incentives for monitoring and evaluation to be done well and for evaluation findings to be actually used.
- (3) Another critical success on the demand side is having a powerful *champion* (minister or senior official) who can drive the institutionalisation of monitoring and evaluation and persuade colleagues of its importance and of the need to devote resources to the project.
- (4) The importance of starting the process with a *diagnosis* of existing monitoring and evaluation.
- (5) A noteworthy feature of successful government evaluation systems is the “*stewardship* of this process by a capable ministry”, often the Finance ministry.
- (6) The importance of building reliable ministry *data systems*.
- (7) The *danger of over-engineering a monitoring and evaluation system* which he notes is a common mistake.
- (8) The importance of the objective of *utilisation* of monitoring and evaluation findings which is the “yardstick of success”.
- (9) The *limitations of a reliance on laws, decrees and regulations* as a means of achieving success.
- (10) The importance of *structural arrangements* (e.g., who commissions the evaluations and who does them) in ensuring objectivity and quality.
- (11) Finally, echoing Boyle and Lemaire, he emphasises that “countries which have built a government monitoring and evaluation system have found that it is a *long-haul effort requiring patience and persistence*”. He further notes that the development of such systems is not a linear process and that many government systems have developed in an incremental and even piecemeal manner (citing Australia, Chile, Ireland and the United States as examples of this pattern). He also emphasises the value of regular evaluation of an monitoring and evaluation system to find out what is working, what is not and why not.

Some of the above lessons are mirrored in the “overarching” recommendations of *McDonald, Rogers and Kefford* (2003), based on their work in helping to develop evaluation capacity over a five-year period at the level of a public sector organisation in *Australia*. Some of the *recommendations* made are as follows:

- *Think big, but start small-stage*, trial and grow evaluation;
- *Address both supply and demand*. Echoing Boyle and Lemaire, they contend that “building an effective evaluation capability requires balancing supply and demand and iteratively raising the level of each”;
- *Work top-down* (i.e., with senior managers) *and bottom-up* (i.e., with project staff and managers) to meet the evaluation needs of both;
- Develop a *common evaluation framework*;
- *Build a knowledge about what works* in evaluation in the specific context in question (an agency in the instance reviewed); and

- Systematically and visibly *evaluate each stage* (this is also emphasised by Mackay, as discussed above).

A recent conference presentation by Lahey (2007) identifies *twelve critical success factors* for an M&E system, building on the four building blocks of vision, an enabling environment, and the infrastructures to supply and to demand and use monitoring and evaluation information outlined earlier. It is worth reproducing the twelve critical success factors, many of which overlap with the lessons from the literature summarised above and with some of the performance indicators developed by MacKay and presented above in table 5.

Table 6. Twelve Critical Success Factors for an M&E System

Factor	Key Considerations
Drivers	Clarify what is driving the need for M&E (e.g. public sector reforms)
Uses	Clarify how M&E information will be used, and, by whom.
Leadership	Leadership must be supportive. Is there a 'champion'?
Commitment	There needs to be a long-term commitment to building a well-functioning M&E system.
Resources	Sources of funding need to be identified to both start up and then sustain M&E capacity.
Accountability	Roles and responsibilities for the operation of the M&E system need to be clearly identified.
Technical capacity	Critical are: reliable data systems; credibility of information being reported; adequate analytical capacity.
Infrastructure to 'supply' M&E information	Key components: Policies and standards to guide M&E practices, Resources (€ and HR) dedicated to both evaluation and performance measurement. Trained and experienced human resources.
Infrastructure to 'demand' M&E Information	Key components: Clear understanding of who are the users and uses of M&E information. Are there incentives (for institutions and individuals) to drive the M&E system?
Oversight	Implementation of M&E is generally long-term and iterative. Will the system be monitored? Adjusted as needed? Sanctions for non-compliance?
Values and Ethics	To be effective, use of M&E must evolve into the culture of organisations. Is there an enabling environment in organisations? Across the system?
Sustainability	What assurance that M&E system can and will be maintained?

Source: Robert Lahey, REL Solutions INC. (2007)

Finally, the factors that support an evaluation function are often referred to in evaluation guidelines or statements of evaluation policy where these exist. For example, the *Good Practice Guidelines for the Management of the Evaluation Function* prepared by the Working Group on Guidelines in 2000 (a European Commission working group chaired by Directorate-General for the Budget) contains four good practice guidelines:

- Under Guideline A (Integration of the evaluation function in the overall management of the DG), it is emphasised that the *evaluation function* should be *clearly distinguished* from the functions of control, audit and inspection and should have close links with policy services.
- Guideline B is concerned with resources for evaluation and specifies that the *evaluation function should have sufficient and appropriate resources* in terms of funds, staff and skills to discharge its responsibilities.
- Guideline C "Quality for evaluation" states that both *evaluation processes and reports should be governed by standards* designed to ensure their quality in terms of relevance, rigour and timeliness.
- Guideline D "External Communication" indicates that, as a general rule, *evaluation reports should be made publicly available*.

The review of *lessons in the building of evaluation capacity* demonstrates that again a strong degree of commonality in the various contributions reviewed is apparent. A couple of issues that are emphasised repeatedly are worthy of note:

- Most contributors highlight the *critical importance of stimulating and sustaining effective demand for evaluation* and the merits of demand “drivers” such as incentives and evaluation champions;
- Another common theme is the *need to work in parallel on the demand- and supply-sides*;
- The importance of *adequate evaluation resources* including data, standards and staffing/expertise;
- The way in which the evaluation function is organised (its *architecture*) and the links between evaluation and other functions is important;
- A final, more general point is the need for *commitment and leadership* to sustain the long-term effort that the building of evaluation capacity demands.

Many of these critical success factors relate back to the key dimensions of demand, supply, architecture and institutionalisation, but also emphasise other factors such as the importance of commitment and leadership in the form of evaluation champions.

3.5. Literature Review: Summary and Conclusions

The literature review has illuminated many of the issues relevant to the development of an evaluation capacity framework that could be applied to Structural and Cohesion Fund administrations in the EU Member States. An encouraging feature is the strong degree of consensus that emerges from the literature as to the key dimensions that underpin evaluation capacity generally. Put another way, there is broad agreement on the strategic issues that governments must address if evaluation capacity is to be nurtured and sustained. Most analyses emphasise the dimensions of evaluation demand, supply and the architecture or organisation of the evaluation function. Another critical and overarching factor is the importance of broader environmental factors that reflect the degree to which a government system or administrative culture is “enabling” of or conducive to evaluation (what Weiss refers to as the “openness of the system to evidence and argument” and Lahey as an “enabling environment”).

Below the level of these dimensions and consistent with them, a range of more detailed criteria and indicators of evaluation capacity can be observed from the literature. All told, on the basis of the review and analysis here and building on the strong degree of consensus in the literature, it is feasible to construct an evaluation capacity framework comprising a hierarchy of broadly internally-consistent dimensions, criteria and indicators. This is taken forward in the next chapter. However, as noted earlier, the frameworks and approaches considered here generally refer to evaluation capacity at governmental or societal level and we need to take account of a number of factors specific to the EU evaluation context in developing our framework.

4. EVALUATION CAPACITY DEVELOPMENT BENCHMARKING FRAMEWORK

In this chapter the framework for benchmarking and analysis of evaluation capacity in the Structural and Cohesion Fund administrations of the EU Member States is presented. The framework is strongly influenced by the literature review presented in the preceding chapter. However, as emphasised earlier, none of the frameworks or approaches found in the literature can be applied unaltered to the EU Structural and Cohesion Funds evaluation sphere; account needs to be taken of contextual factors and features specific to the EU evaluation setting. Some of these features are reviewed in the section immediately below before presenting the definition and components of the evaluation capacity development framework. In the first instance, a number of salient features or characteristics of the evaluation of programmes funded by the EU Structural and Cohesion Funds are summarised. Secondly, the evaluation requirements embodied in the Council Regulations for the 2007 to 2013 period and the challenges these pose for EU Member State administrations are considered.

4.1. Context for Evaluation of EU Structural and Cohesion Funds

4.1.1. Some Features of the EU Structural and Cohesion Funds Evaluation Model

A number of papers have highlighted *important features of the evaluation of programmes funded by the EU Structural and Cohesion Funds* (for example, Lang, 2003, Toulemonde and Bjornkilde, 2003, Casavola and Tagle, 2003, Hegarty, 2003 and Polverari and Bachtler, 2004). A number of specific points of emphasis are worth noting in the context of this study:

- In terms of the institutional setting, *responsibility for evaluation of Structural and Cohesion Fund programmes* - as with the other programme management functions - *rests largely with the Member State and regional authorities* (with the exception of ex post evaluation which is a Commission responsibility). While the Commission is an active partner in the process its role is mainly one of ensuring adherence to the regulatory requirements, issuing guidance and generally promoting good practice.
- Many commentators highlight the innate *complexity of the evaluation* of Structural and Cohesion Fund programmes. This reflects, inter alia, the increasingly intangible and complex nature of the objectives of regional and structural policy itself (Lang, 2003 and Polverari and Bachtler, 2004) and the “ambitious” nature of the EU evaluation model, given the need to address the needs of several levels of government (EU, national and regional) the emphasis on indicators and the combination of learning (formative) and accountability (summative) dimensions (Toulemonde and Bjornkilde, 2003).
- This complex and challenging evaluation project can be such as to lead to *disappointing outcomes in some instances*. For example, Casavola and Tagle (2003) note that some features of the EU evaluation framework, such as the imperative to evaluate programmes across the board, “seriously impair the possibility to obtain good evaluations” and of a danger that evaluation can be reduced to a compliance-focused, ritualistic exercise.
- On a more positive note, a number of commentators have argued that the nature of the EU evaluation model and the *requirement* to meet the demands of the regulations have acted as an *impetus to the development of capacity* at the Member State level (Hegarty, 2003; Toulemonde and Bjornkilde, 2003).

4.1.2. Future Challenges

Looking ahead to the *2007 to 2013 programming period*, a number of *specific challenges* in the evaluation area stem from the provisions of Council Regulation (EC) 1083/2006, particularly as regards on-going evaluation. The most significant change - compared to the 2000 to 2006 period - is the scrapping of the requirement to conduct mid-term evaluations of programmes by a fixed regulatory deadline. Drawing on the experience of the mid-term evaluation exercise in the previous period, the General Regulation embodies a shift towards “a more flexible, demand-driven approach to evaluation during the programming period” in the form of on-going evaluation.²¹ Based on the provisions of the Regulation, the Commission’s Working Paper envisages that on-going evaluations will be carried out on the basis of evaluation plans to be prepared at Member State and programme/regional level. The ambition embodied in the Regulation and the Commission’s Working Paper is one where evaluation is used as an integrated management tool in programme implementation. These requirements will “test” the capacity of Member States administrations in a number of ways:

- In a general sense, the *on-going evaluation process will have to be managed in an efficient and timely manner* if, as intended, evaluations are to be successfully integrated into programme management;
- The need to *anticipate future evaluation requirements* in the form of an evaluation plan constitutes a new evaluation management task; and
- The successful undertaking of performance-related evaluations requires that *indicator and monitoring systems are well designed* such that they are capable of providing reliable “early-warning signals” to managing authorities (e.g., where programmes are running behind target) in addition to providing the essential data inputs to evaluation projects.

In summary, these new tasks will pose serious questions of the evaluation capacity of the EU Member State administrations. Whether the administrations are equipped to meet these challenges is a central focus of this study.

4.2. Working Definition of Evaluation Capacity

A number of definitions of evaluation capacity were reviewed in section 3.1 of this Report. These definitions tend to be quite generic in nature referring to the underlying processes that support the creation and sustenance of evaluation practice and utilisation in a public sector context. Building on these underlying themes, but taking account of the context factors discussed above, the following **working definition of evaluation capacity** is proposed for the purposes of this project:

The ability of Structural and Cohesion Fund administrations to effectively organise and utilise timely, high-quality evaluations of EU-funded programmes.

A number of *features of the definition* proposed above can be noted:

- The definition is *specific* to that part of Member State administrations responsible for the Structural and Cohesion Funds;
- In the light of the requirements for the 2007 to 2013 period discussed earlier, the definition has an operational edge, focusing on administrations’ *organisational capability* to produce evaluations in a timely manner;

²¹ European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf

- *Utilisation* is emphasised - echoing a leading commentator in the field, we take the view that utilisation is the “yardstick of success” (Mackay, 2006).

Building on this definition of evaluation capacity and having regard to the various frameworks surveyed in the literature review, we propose a framework to benchmark and analyse evaluation capacity in the Structural and Cohesion Fund administrations of the EU Member States.

4.3. Structure of Evaluation Capacity Benchmarking Framework

Our evaluation capacity benchmarking framework is structured around four core dimensions as follows:

- (1) *Demand*;
- (2) *Supply*;
- (3) A third dimension concerned with issues relating to the *dissemination and utilisation* of evaluation outputs;
- (4) A fourth *institutionalisation* dimension which aims to capture whether the wider (i.e., non-EU) administrative environment or culture is conducive to, or enabling of, evaluation (a secondary, but nonetheless important, focus is on whether EU evaluation requirements influence Member State evaluation practice).

In broad terms, the approach is in keeping with the demand/supply/architecture/institutionalisation approach articulated in the EVALSED, an approach that itself is similar to some of the other frameworks reviewed earlier in section 3.2 (namely those of the World Bank, Boyle and Lemaire and Lahey.) For example, three of the four EVALSED dimensions (demand, supply and institutionalisation) are explicitly included in our framework. The main *points of difference* are as follows:

- We include issues concerned with the architecture or organisation of the Structural and Cohesion Fund evaluation system under the demand dimension rather than as a separate dimension. As noted earlier, there is a strong overlap between these dimensions generally: evaluation demand is strongly influenced by way in which evaluation systems are organised (the corollary also applies, in that the demand for evaluation influences the organisation of the evaluation system). In the case of the EU Structural and Cohesion Fund administrations, this interdependence is particularly marked. Thus architecture issues are assessed as a sub-set of the demand-side (for example, Evaluation Units – a structure that is a common feature of EU Structural and Cohesion Fund administrations – are often strong drivers of demand).²²
- In the literature review, we noted that “institutionalisation” sometimes features as a specific dimension concerned with utilisation issues or, in other instances, as an overarching notion concerned with the extent to which evaluation is integral to the policy-making process (see section 3.2.5). Our approach is to include a specific dimension in the framework to assess issues around the dissemination and utilisation of evaluation outputs within the EU Structural and Cohesion Funds sphere. This “internal” institutionalisation dimension (i.e., internal to the EU Structural and Cohesion Funds administrations) is complemented by a broader, external institutional dimension that aims to capture the influence of wider Member State background factors.
- The latter takes the form of a fourth “catch-all” dimension concerning wider environmental or background conditions in recognition of the reality that the evaluation capacity of Structural and

²² An evaluation unit can also be a source of evaluation supply, an issue that is discussed later.

Cohesion Fund administrations is influenced by the culture of the wider administration of which it forms part. The importance of broader, cultural factors in promoting evaluation capacity is a prominent theme in the literature. The institutionalisation dimension aims to assess the extent to which the wider administrative and cultural setting is “enabling” of or conducive to evaluation practice. While, as noted above, it is acknowledged that EU evaluation requirements have supported the development of evaluation capacity in some Member States, it is unlikely that a strong EU evaluation capacity could be sustained where the wider governance environment was not conducive to evaluation practice.

As outlined above, the capacity framework proposed is built around the above four *dimensions*. At the next level, each dimension is broken-down into a series of *criteria*, in turn decomposed into *sub-criteria* and then broken down into *indicators* at the lowest level. This is broadly in line with the manner in which the literature review was organised in the previous chapter. However, it is important to bear in mind that the hierarchy and terminology used is specific to this framework; from the literature review it is clear that there is no structure or vocabulary common to the various evaluation capacity development frameworks reviewed (with, for example, terms such as “criteria” and “indicators” often used interchangeably).

In the sub-sections below, each of the four dimensions is introduced before the presentation of the full framework.

4.3.1. Demand Dimension

Under the demand dimension, four criteria are proposed as follows:

- (i) **Integration, location and organisation of the evaluation function.** As argued above, issues concerned with the organisation or the architecture of the evaluation system are addressed under the demand dimension. The architecture of evaluation systems is an important theme in the evaluation capacity development literature. For example, it is the subject of one of the four guidelines in the European Commission’s Guidelines for the Management of the Evaluation Function (2000) reviewed in section 3.4. The “anchoring” of evaluation capacity is also emphasised by Boyle and Lemaire (1999).
- (ii) **Evaluation resources (financial and human).** Mackay (2006) refers to the need to devote resources to evaluation. “Resourcing” is one of the critical success factors identified in the framework developed by Lahey (2007). In some frameworks (such as that of the World Bank as articulated, for instance, in Guerrero 1999) resources are presented as a supply-side aspect. However, in the EU Structural and Cohesion Funds context the volume of financial and human resources devoted to the evaluation function is clearly a demand-side variable. We deal with some other evaluation resources such as the availability of skills, expertise and training options under the supply dimension.
- (iii) **Efficiency of evaluation management.** This criterion is included in consideration of the requirements and expectations embodied in the Cohesion Policy Regulations for the 2007-2013 period discussed above at section 4.1. As such, the rationale for its inclusion is quite specific to the EU context. Under this criterion, a focus is on a number of indicators aimed at assessing whether Member State administrations are likely to be in a position to use evaluation as intended, i.e. as an integrated and flexible management tool.
- (iv) **Quality of monitoring systems.** This is emphasised in almost all of the evaluation capacity development frameworks reviewed in chapter 3. It goes without saying that high quality monitoring systems that are capable of providing reliable and timely information are essential. In the international literature, monitoring systems are generally presented as a supply-side consideration (from the perspective of an evaluation practitioner, the monitoring system is a supply variable). However, in the EU context, the monitoring systems for Structural and

Cohesion Fund programmes are clearly the responsibility of the managing authorities and other arms of the Structural and Cohesion Fund administrations. On this basis, they are included under the demand dimension.

Another demand-side factor mentioned in the literature relates to the existence of regulatory or legal requirements to conduct evaluation. However, for Structural and Cohesion Fund programmes, this is a “given”, in that EU Member States are obliged to implement the evaluation requirements of the Structural and Cohesion Fund Regulations. The inclusion of such a criterion in the framework would add no explanatory value and is therefore omitted.

4.3.2. Supply Dimension

On the supply-side, we focus on two criteria as follows:

- (i) **Quality of external socio-economic data sources.** The importance of reliable data sources was emphasised in the literature and also noted above under the monitoring systems criterion under the demand dimension. In the context of the evaluation of Structural and Cohesion Fund programmes, the availability of reliable external data on key socio-economic variables (such as GDP, employment and unemployment etc.) is essential in view of the wide range of policy objectives that Structural and Cohesion Fund programmes seek to address. However, unlike programme monitoring data, the availability and quality of socio-economic data is clearly not something that is within the control of the Structural and Cohesion Fund administrations. For this reason, it is included here on the supply-side.
- (ii) **Availability and quality of evaluation expertise.** The availability of the necessary skills and expertise to conduct evaluations is a core supply-side component of all the models and frameworks reviewed earlier. Generally speaking the supply of evaluation skills is taken to refer to resources that are external to the administration in consultancy firms, universities etc. However, in some EU Member State Structural Fund administrations, internal evaluation units attached to the managing authorities represent an important source of evaluation supply (while also operating on the demand-side in terms of organising and commissioning evaluations). We propose to take account of this complexity (and interdependence between the demand- and supply-sides) in the supply dimension indicators.

As noted above, some criteria that are traditionally included on the supply-side in evaluation capacity development frameworks, such as the financial resources devoted to evaluation and monitoring system quality, are included on the demand-side in this framework. This reflects the reality that responsibility in each case lies with the Structural and Cohesion Fund administration and is thus better addressed as a demand-side issue.

4.3.3. Dissemination and Utilisation Dimension

Under this dimension, we focus specifically on the two criteria of **dissemination and utilisation**. The rationale for the inclusion of these criteria is straightforward. Unless evaluations reach the people who matter (policymakers, decision-makers and civil society generally) they cannot be expected to have influence. This can be seen as a necessary, though not sufficient, utilisation condition.

Of course, the process through which evaluations influence programme design and management and the policy process generally is a complicated one. The influence of evaluation is often of a long-term nature and not something that can always be distinguished from other factors such as policy developments at national or EU level. Moreover, it must be understood that there can be no supposition that EU Member States are “expected” or obliged to implement evaluation

recommendations. Nevertheless, it is worthwhile to explore whether some basic utilisation conditions are in place - such as the existence of systems or mechanisms for evaluations to be formally considered by monitoring committees and other relevant authorities and for agreed evaluation recommendations to be tracked. We also attempt to assess the influence of evaluations in the 2000-2006 period and of the more recent ex ante evaluations of 2007 - 2013 programmes. These and other aspects are the focus of the “utilisation” criterion.

4.3.4. Institutionalisation Dimension

Finally, under the fourth institutionalisation dimension a range of more general issues related to evaluation capacity are explored. The focus here is on wider factors, external to the Structural and Cohesion Fund administrations, which capture the extent to which evaluation is embedded in the wider Member State governance system. As argued above, and emphasised by a number of contributions reviewed in the literature survey, these background conditions are critical to the development and sustainability of evaluation capacity. These conditions will be assessed under a number of indicators including the role of evaluation in the national budgetary and policy-making processes and the existence of mandatory national evaluation requirements. However, we also recognise that Structural and Cohesion Fund evaluation requirements can in themselves promote national evaluation capacity and this is reflected in the indicators.

4.4. Evaluation Capacity Development Benchmarking Framework

The detailed components (dimensions, criteria, sub-criteria and indicators) of the evaluation capacity benchmarking framework are presented in tabular form below. As noted above, each dimension is broken-down into a series of criteria and sub-criteria and, in turn, indicators. (The last two columns in the table specify the relationship between benchmarking indicators and respective questions in the survey questionnaires which are used in the fieldwork stage, as explained in section 2.2.1).

**Evaluation Capacity Development Benchmarking Framework:
Dimensions, Criteria, Sub-criteria and Indicators**

Dimension	Criterion	Sub-criteria	Indicators	Respective Questions in the Benchmarking Questionnaires ²³	
				For Administrators	For Evaluation market participants
DEMAND SIDE	Architecture (Integration, location and organisation of Structural and Cohesion Fund evaluation function)	Evaluation Responsibility	1. Location of responsibility for evaluation function within the managing authority/authorities 2. Existence of dedicated evaluation function (Evaluation Units / national and programme level) 3. Existence of an evaluation strategy or policy statement that delineates evaluation roles and responsibilities and the role of evaluation in programme management	A1 A2 A3	A1
		Coordination	4. Existence of formal mechanisms for evaluation coordination across the programmes	A4	
		Links with other functions	5. Linkages between evaluation function and other related functions, such as monitoring and programming	(Case studies)	
	Evaluation Resources	Financial	6. Evaluation budgets (euros)	A5; A7	Not applicable
		Staffing	7. Number of evaluation staff (national and programme) 8. Evaluation experience and skills of evaluation staff ○ Evaluation commissioning experience ○ Evaluation practice experience	A6; A7 A8-A12; B5	A2

²³ Questionnaires are enclosed in Annex No.2.

Dimension	Criterion	Sub-criteria	Indicators	Respective Questions in the Benchmarking Questionnaires ²³	
				For Administrators	For Evaluation market participants
		Evaluation flow and coverage	9. Number of evaluations carried out / time period	A12; A13	Not applicable
	Efficiency of evaluation management	Planning	10. Existence of evaluation plans at national and programme level	A14	A3
		Terms of Reference	11. Quality and clarity of evaluation Terms of Reference (TOR) <ul style="list-style-type: none"> o Existence of core or mandatory evaluation questions 	(Case studies)	(Case studies)
		Procurement	12. Average interval between agreement on TOR and appointment of evaluator	A15	Not applicable
		Evaluation Steering Committees	13. Existence of Evaluation Steering Committees (national and programme level)	A16	Not applicable
		Quality Assessment	14. Existence of approved set of quality standards	A17	Not applicable
	Quality of Monitoring Systems	Reliability	15. Quality and reliability of indicator data	(Case studies)	(Case studies)
		Timeliness	16. Average interval between year end and availability of OP indicator data for year in question	A18	A3
SUPPLY SIDE	Quality of socio-economic data resources	Availability	17. Availability of key socio-economic indicator data (GDP, employment, unemployment, R&D investment) at regional (NUTS II) level	B1	B1
		Timeliness	18. Latest year for which data is currently available	(Case studies)	(Case studies)
	Availability and quality of evaluation expertise	Availability of evaluation supply on domestic market	19. Number of firms active in the market 20. Involvement of universities in evaluation practice 21. Existence of internal evaluation unit capacity 22. Availability of evaluation training/education options	B2; B3 B4 B5; B6 Not applicable	Not applicable B2

Dimension	Criterion	Sub-criteria	Indicators	Respective Questions in the Benchmarking Questionnaires ²³	
				For Administrators	For Evaluation market participants
		Thematic and methodological expertise	23. Breadth and quality of expertise (areas where expertise is particularly strong or deficient: e.g., macroeconomic modelling, horizontal principles, infrastructure, Human Resources etc.) 24. Quality of evaluation reports	B8; B9 B7	Not applicable
DISSEMINATION AND UTILISATION	Dissemination	Ease of access to evaluation outputs	25. Public access to evaluation reports	C1; C2	C1; C2
	Utilisation	Procedures for addressing evaluation and follow-up procedures	26. Existence of formal monitoring committee or managing authority responsibilities 27. Existence of follow-up procedures where evaluation recommendations are agreed	C3 C4	Not applicable C3
		Impacts on programme design and implementation	28. Degree of influence of ex ante evaluation of 2007-2013 period on: <ul style="list-style-type: none"> o Strategy formulation and programme design o Allocation of resources o Quantification of objectives and indicator selection 29. Degree of influence of evaluation on ongoing programme management and implementation (2000 to 2006 period)	C5-C7 C5 C6 C7 (Case studies)	C4-C7 C5 C6 C7 (Case studies)
INSTITUTIONALISATION OF	An enabling governance environment?	Role of Evaluation in State Budget/degree of integration of evaluation into budget process	30. Degree to which evaluations represent routine part of budgetary process	D1	D1
	(Degree of institutionalisation of evaluation in the wider governance system)	Role of evaluation in sectoral strategy/policymaking	31. Extent to which evaluations are routinely used in strategy and policy formulation at sectoral/ministry level	D2	D2

Dimension	Criterion	Sub-criteria	Indicators	Respective Questions in the Benchmarking Questionnaires ²³	
				For Administrators	For Evaluation market participants
		Existence of mandatory evaluation requirements	32. Existence of statutory evaluation requirements	D3	Not applicable
		Public management reforms	33. Existence of ongoing programme of public management reform	(Case studies)	Not applicable
		Public policy research base	34. Existence of public policy research bodies or think-tanks	D4	Not applicable
		Existence of an evaluation society or network	35. Membership size 36. Frequency of meetings/conferences	Not applicable	(Case studies) D3-D4
		Influence of EU evaluation requirements	37. Degree of influence of EU Structural Funds evaluation requirements on Member State evaluation practice	D5; D6	D5; D6
		Role of civil society and other players	38. Extent to which evaluation findings are drawn on by civil society and social partners	D7	D7

In total, the evaluation capacity development benchmarking framework comprises *4 dimensions, broken-down into 9 criteria, 28 sub-criteria and 38 indicators in total*. We have deliberately erred on the side of comprehensiveness and included most, if not all, of the factors identified as relevant in the literature as well as some additional aspects specific to the Structural and Cohesion Funds evaluation context. We aimed to produce a robust, comprehensive framework even if hard information on some of the evaluation capacity criteria and indicators proved difficult to collect at the fieldwork stage of the project. In other words, we did not allow considerations around the feasibility of data collection to restrict unduly the choice of indicators at the preliminary stage. The fieldwork is intended to test this framework and provide basis for any possible modifications of the framework. However, at the initial stage, it was important to include as wide a range of issues as possible.

The proposed evaluation capacity *indicators comprise both fact- and opinion-based* (or, looked at another way, quantitative and qualitative respectively) *variables*. Both types of information are essential to form a rounded view of evaluation capacity in Member States' administrations.

Looking at the different dimensions, most of the *indicators under the demand* dimension relate to factual or observable variables (except perhaps for the indicator relating to the quality of evaluation terms of reference which is more a matter of opinion). Similarly, the *indicators under the supply-side* dimension (relating to the quality of socio-economic data and evaluation expertise) generally relate to variables that can be directly observed even if the information in question might not be readily accessible or might take some time to collect and/or verify. Under the *dissemination dimension*, the indicators relating to the accessibility of evaluation reports and follow-up procedures lend themselves to a binary, yes/no type response. However, the *utilisation indicators* relate more to matters of judgement on which different opinions are possible. Under the *institutional dimension*, the indicators are a mix of both observables and opinions with some (e.g., the indicator relating to the role of evaluation in the State Budget) lying somewhere between the two types.

Following development of this draft framework, the fieldwork aimed to test it and to establish how well it served the purpose of assessing and analysing evaluation capacity development in different EU Member States was carried out. The results of the fieldwork are discussed in the following chapter.

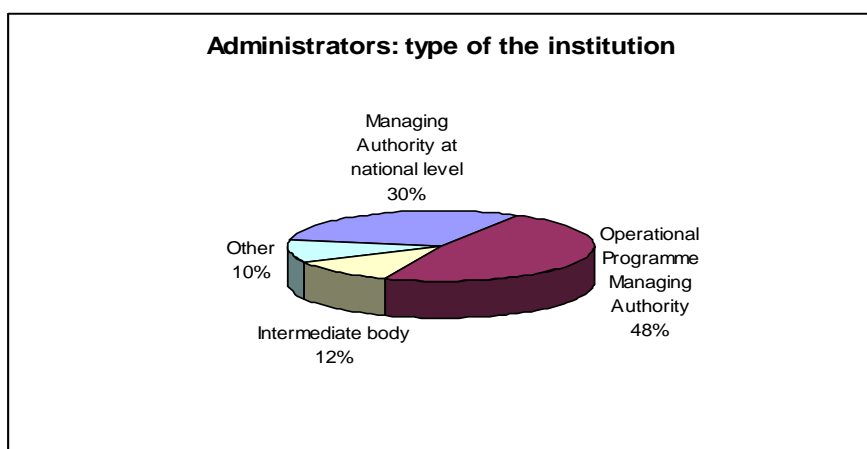
5. RESULTS OF THE FIELDWORK

This chapter presents the results of the fieldwork carried out to apply and test the benchmarking framework. The analysis of evaluation capacity development in the EU Member States draws primarily on survey evidence; the case study findings are not widely generalisable given that only 4 countries were studied. However, they are used to complement, illustrate and inform survey results where possible. The combination of survey results and case study findings provides insights concerning evaluation capacity development within Member State administrations and the basis for judging the application of the benchmarking framework and suggesting revisions.

5.1. Structural and Cohesion Fund Evaluation Capacity Development in the EU Member States

5.1.1. Survey Statistics

As outlined in section 2.2.1, a cross-national online survey on Structural and Cohesion Fund evaluation capacity was conducted across the 27 EU Member States. Links to online questionnaires were sent to 446 respondents from Structural and Cohesion Fund *administrations*. 128 responses covering all EU Member States were received. The majority of those who participated in the survey work in Managing Authorities, at either national or programme level, in positions of Heads of the Department or Senior Officers.



$n=128$

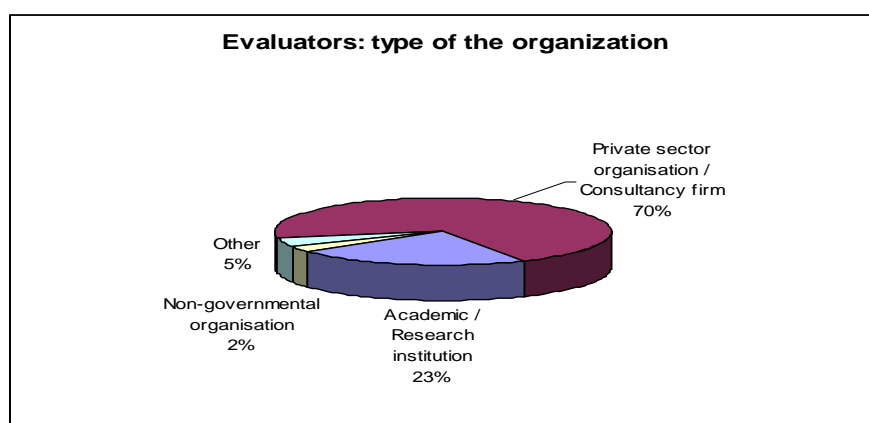
The overall response rate achieved for administrations is 29 percent but with significant Member State variation.²⁴ As table 7 below shows, the highest response rates – over 50 percent – were received from Bulgaria, Poland, Estonia, Romania and Lithuania. In contrast the response rate from countries often identified in literature as more advanced in Structural Fund evaluation practice (i.e., Northern countries) was much lower. This can be seen as indication that the topic is more pressing in the new Member States.

²⁴ It is noteworthy that the response rate also varied according to a question in the questionnaire.

Table 7. Survey response rate: Structural and Cohesion Fund administrations

EU Member State	Number of questionnaire recipients (Structural and Cohesion Fund administrations)	Number of responses received	Response rate %
Austria	14	2	14
Belgium	11	2	18
Bulgaria	12	9	75
Cyprus	5	1	20
Czech Republic	31	11	36
Denmark	6	2	33
Estonia	9	5	56
Finland	9	1	11
France	38	6	16
Germany	22	5 ²⁵	23
Greece	37	6	16
Hungary	11	5	46
Ireland	7	1 ²⁶	14
Italy	29	2	7
Latvia	15	7	47
Lithuania	15	8	53
Luxembourg	5	2	40
Malta	4	1	25
Netherlands	16	4	25
Poland	16	11	69
Portugal	25	2	8
Romania	15	8	53
Slovakia	19	7	37
Slovenia	6	2	33
Spain	23	6 ²⁷	26
Sweden	17	4	24
United Kingdom	29	8	28
Total	446	128	29

For *evaluation market participants*, survey invitations were sent out to 509 respondents and a total of 125 responses were received.²⁸ The majority of respondents work in the private sector.



n=125

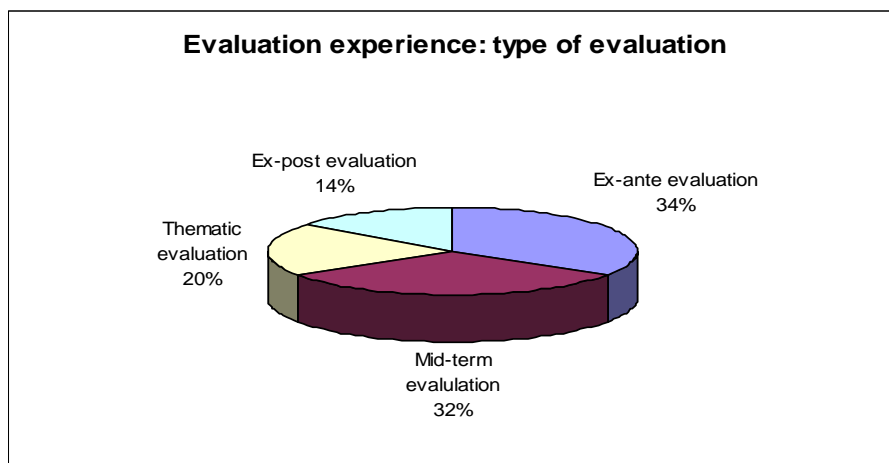
²⁵ No responses from national level institutions.

²⁶ No responses from national level institutions.

²⁷ No responses from national level institutions.

²⁸ Not all respondents provided their answers to each question of the questionnaire.

Generally speaking, their most recent evaluation experience is in the field of ex-ante or mid-term evaluation with a majority of respondents indicating that their most recent evaluation project was finalised in the last two years.



The questionnaire return rate for this group of respondents is 25 percent, a little lower than for the Structural and Cohesion Fund administrator group, but again with significant variation across the Member States. For some countries only one response was received. The response rate is highest for Estonia, Ireland and Netherlands.

Table 8. Survey response rate: evaluation market participants

EU Member State	Number of questionnaire recipients (Evaluation market participants)	Number of responses received	Response rate %
Austria	11	1	9
Belgium	17	6	35
Bulgaria	4	1	25
Cyprus	2	1	50
Czech Republic	49	8	16
Denmark	6	1	17
Estonia	14	7	50
Finland	28	9	32
France	23	5	22
Germany	35	8	23
Greece	32	6	19
Hungary	21	7	33
Ireland	11	6	55
Italy	21	2	10
Latvia	12	5	42
Lithuania	14	3	21
Luxembourg	4	2	50
Malta	5	1	20
Netherlands	11	6	55
Poland	37	5	14
Portugal	43	7	16
Romania	14	2	14
Slovakia	6	1	17
Slovenia	8	2	25
Spain	15	3	20
Sweden	18	4	22

EU Member State	Number of questionnaire recipients (Evaluation market participants)	Number of responses received	Response rate %
United Kingdom	48	16	33
Total	<i>509</i>	<i>125</i>	25

The variance in country response rates hinders a representative cross-comparison of Member States, especially when it comes to opinion-based questions. Then again, as already mentioned, the primary emphasis of this project was not on conducting a representative benchmarking exercise, but on testing the proposed evaluation capacity benchmarking framework. Moreover, “factual” or quantitative data can still be considered and compared. However, when contradictory or inconsistent responses to such questions are received (even from the same institutions, as was the case in the majority of cases), it is impossible to draw conclusions.

Overall, despite a limited response rate and its variance across the EU Member States, the survey in combination with case studies of four selected Member States allow the identification of certain general tendencies in Structural Fund evaluation capacity development and, more importantly, provides insights as regards the application of the proposed benchmarking indicators. The sections below discuss fieldwork results according to the four dimensions of the evaluation capacity benchmarking framework.

5.1.2. The Demand Side Dimension

The assessment of the demand side is based mainly on answers provided by the Member State administrators responsible for evaluation of the EU Structural and Cohesion Fund programmes. It is based on the following criteria: institutional architecture, resources allocated to evaluation, efficiency of evaluation management and quality of monitoring systems.

Institutional Architecture

As a first step, the survey attempted to establish the *institution responsible* for Structural and Cohesion Fund evaluation at national level. The results are presented in table 9 below.

Table 9. Institutions responsible for Structural and Cohesion Fund evaluation at national level

EU Member State	Institution responsible for Structural and Cohesion Fund evaluation at national level
Austria	Coordination department in the Federal Chancellery (Bundeskanzleramt Abteilung IV/4; ÖROK-Geschäftsstelle)
Belgium	n.a.
Bulgaria	Management of EU Funds Directorate, Ministry of Finance
Cyprus	Independent Unit within Managing Authority
Czech Republic	Evaluation and analysis unit at the Department of European Funds of the Ministry for Regional Development
Denmark	National Agency for Enterprise and Construction (NAEC)
Estonia	Ministry of Finance
Finland	Structural Funds Unit at the Ministry of Labour (for ESF only)
France	<ul style="list-style-type: none"> ▪ European unit of the DIACT for Regional competitiveness and employment objective ▪ Secretariat d'Etat a l'outre-mer for Convergence objective
Germany	<ul style="list-style-type: none"> ▪ Federal Ministry of Labour and Social Affairs, Referat VIa3 (for ESF) ▪ Ministry of Finance, Department Managing Authority
Greece	Ministry of Economy and Finance, Special Service for Strategy, Planning, & Evaluation of Development Programmes

EU Member State	Institution responsible for Structural and Cohesion Fund evaluation at national level
Hungary	Unit of Evaluation and Methodology at the National Development Agency
Ireland	Central Evaluation Unit for the Irish NDP/CSF
Italy	<ul style="list-style-type: none"> ▪ Structural Funds Unit, Department for Development Policies, Ministry for Economic Development ▪ Ministry of Labour and Social Policy (for ESF) - in 2007-2013
Latvia	Strategic planning unit in the EU Funds Department of the Ministry of Finance
Lithuania	Evaluation Unit of the EU Cohesion Fund Policy and Structural Assistance Coordination Department of the Ministry of Finance
Luxembourg	Each managing authority is responsible for the evaluation of its own fund.
Malta	Planning and Priorities Coordination Division at the Office of the Prime Minister
Netherlands	<ul style="list-style-type: none"> ▪ Ministry of Economic Affairs ▪ Ministry of Social Affairs and Employment, DG Labour Market and Social Assistance
Poland	National Evaluation Unit at Ministry of Regional Development
Portugal	Observatorio do QCA III (en. CSF III Observatory) operating under the Ministry for Environment, Territorial Planning and Regional Development
Romania	Evaluation Central Unit, Ministry of Economy and Finance
Slovakia	Department for Monitoring and Evaluation at the Ministry of Construction and Regional Development
Slovenia	EU Cohesion Policy Department at the Government Office for Local Self-Government and Regional Policy
Spain	Section of Territorial Programming and Evaluation of European Programmes at the Ministry of Economy and Finance
Sweden	NUTEK (the Swedish Agency for Economic and Regional Growth)
United Kingdom	<ul style="list-style-type: none"> ▪ Department of Communities and Local Government ▪ ESF Evaluation Team, Jobseeker and Skills Division DWP (Evaluation of ESF in England and Gibraltar)

n.a. – not available.²⁹

There are differences apparent in the way EU Member States organise the Structural and Cohesion Fund evaluation function. At least one third of countries have dedicated evaluation units established at national level (see table 10 below). Several countries plan to set up such units in future. Some Member States organise evaluation of ESF-funded programmes separately from those funded by the ERDF. In a number of countries evaluation units are set up also at programme level.³⁰ In other Member States the responsibility for evaluation at programme level is assigned to dedicated staff members or is just one of the functions Structural and Cohesion Fund administrators carry out.

Table 10. Dedicated evaluation units

EU Member State	Dedicated Unit for Evaluation at National Level	Evaluation Unit at Programme Level ³¹
Austria	no	n.a.
Belgium	n.a.	n.a.
Bulgaria	National Evaluation Unit for Phare will overtake functions of national coordination of Structural and Cohesion Fund evaluations	yes

²⁹ "Not available" is given either when information was not provided or when answers were contradictory.

³⁰ The number of evaluation units at programme level and whether they exist for each programme in the country concerned or just for certain programmes could not be established from survey data (as not all programmes were represented).

³¹ Indicated „Yes“, if at least one programme in the country has an evaluation unit.

EU Member State	Dedicated Unit for Evaluation at National Level	Evaluation Unit at Programme Level ³¹
Cyprus	Evaluation Unit	n.a.
Czech Republic	Evaluation and analysis unit at the Department of European Funds of the Ministry for Regional Development	yes
Denmark	n.a.	n.a.
Estonia	no	n.a.
Finland	n.a.	n.a.
France	There is a plan to set up a national committee in charge of the evaluation of ERDF OPs and national development programmes.	yes
Germany	n.a.	yes
Greece	Planning and Evaluation Unit	yes
Hungary	Department of Evaluation and Methodology at National Development Agency	no
Ireland	Central Evaluation Unit for the Irish NDP/CSF	n.a.
Italy	<ul style="list-style-type: none"> ▪ Evaluation Unit of the Department for Development Policies, Ministry for Economic Development ▪ Unit within the DG for Guidance and Training Policies at the Ministry of Labour and Social Policy responsible for the national ESF evaluation. 	n.a.
Latvia	n.a.	n.a.
Lithuania	Evaluation unit at EU Cohesion Policy and Structural Support Coordination Department of the Ministry of Finance	no
Luxembourg	n.a.	n.a.
Malta	no	n.a.
Netherlands	n.a.	n.a.
Poland	National Evaluation Unit at the Ministry of Regional Development.	yes
Portugal	Observatorio do QCA III (<i>en.</i> CSF III Observatory), operating under the Ministry for Environment, Territorial Planning and Regional Development	n.a.
Romania	Evaluation Central Unit at the Ministry of Economy and Finance	yes
Slovakia	no	n.a.
Slovenia	no	n.a.
Spain	n.a.	yes
Sweden	no	n.a.
United Kingdom	no.	Yes, for ESF OP for England

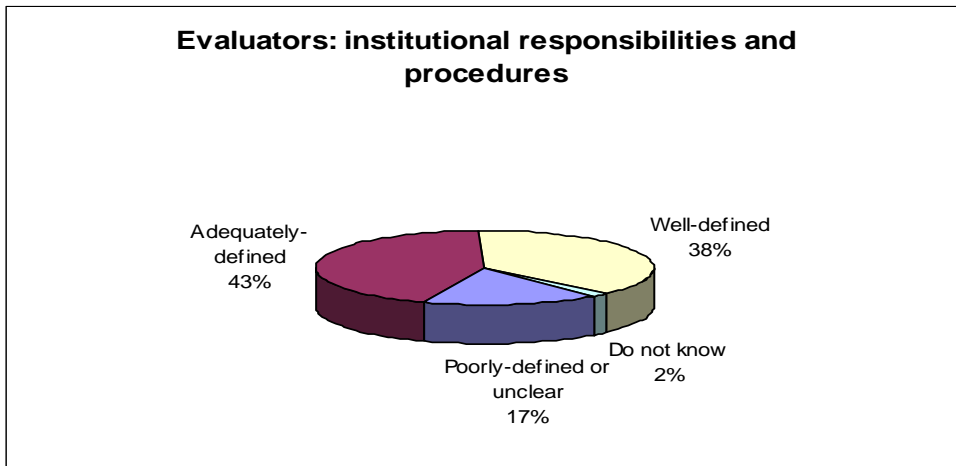
n.a. – not available.³²

An unambiguous answer concerning the existence of a *document setting out institutional responsibilities* for Structural and Cohesion Fund evaluations could not be obtained for approximately half of the EU Member States as responses to this question were too contradictory. The majority of the rest indicated that they had such documents in place or were in the course of preparing them.

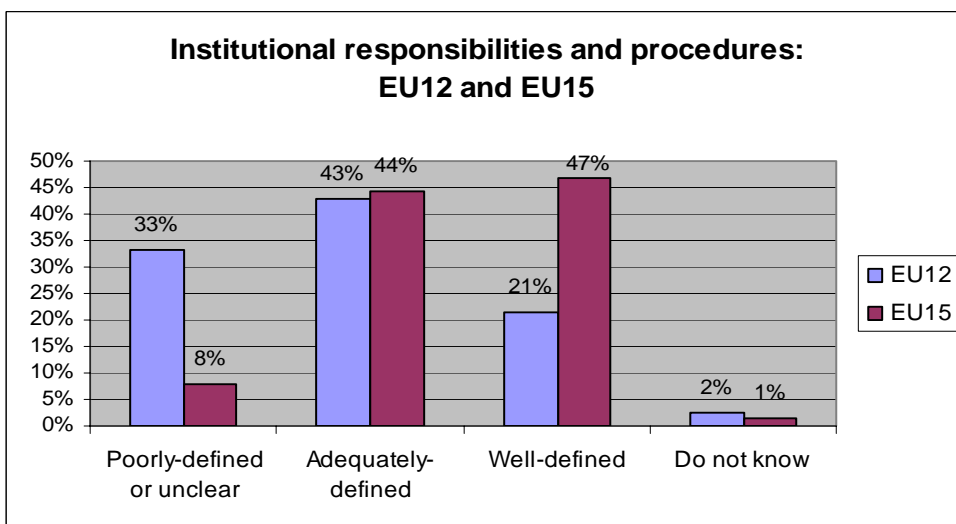
In terms of *formal mechanisms for national coordination* of Structural and Cohesion Fund evaluation, steering groups are the most widely used instrument. Evaluation units also perform coordination functions, but in fewer Member States. Several countries indicated they had no formal mechanisms for coordination.

³² “Not available” is given either when information was not provided or when answers were contradictory.

Evaluation market participants were asked for their views concerning the definition of responsibilities and procedures for the Structural and Cohesion Fund evaluation function in their country. Answers received indicate that views are largely positive: institutional responsibilities and procedures for Structural and Cohesion Fund evaluation function are most often seen as adequately-defined or well-defined. However, it is possible to observe that evaluators from the EU12 are more critical in this respect than their counterparts from the EU15.



n=119



n=119

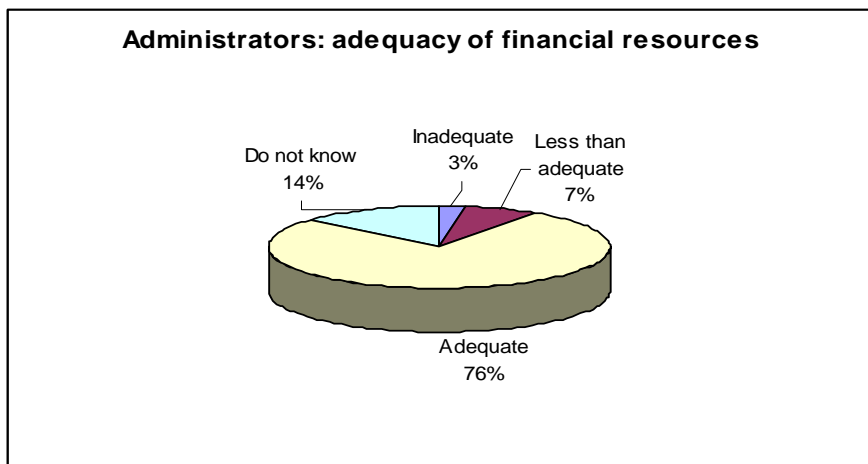
Evaluation Resources

In order to assess the *flow of evaluation activity*, administrators were asked about a number of evaluations undertaken of “their” programme in the 2000-2006 programming period. 39 percent of respondents did not answer this question. Where responses were provided, the number of evaluations specified ranges from zero to fifty. The most frequent responses are that one or two evaluations have been carried out. Half of the respondents indicated that evaluations undertaken included non-mandatory evaluations.

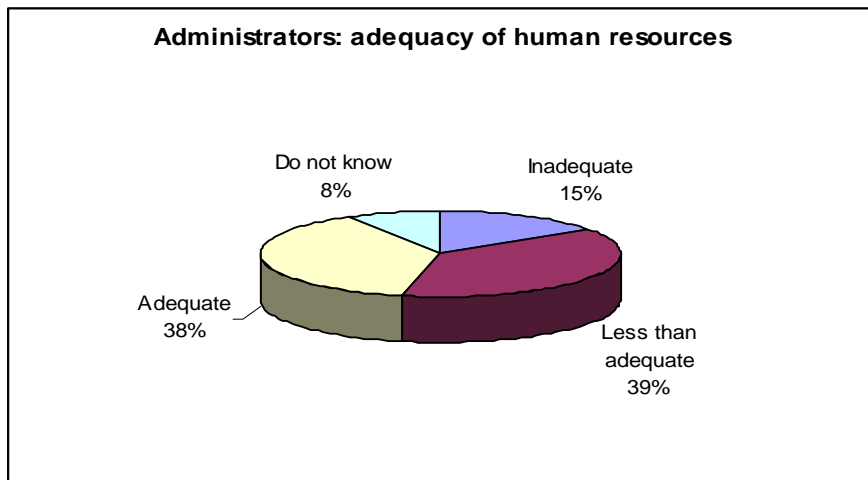
As regards *financial resources* allocated for Structural and Cohesion Fund evaluations, only 16 percent of officials gave figures for the 2000-2006 programming period, with a somewhat higher proportion (i.e. 21 percent) providing estimates for the 2007-2013 period. These figures cannot be relied on without cross-checking them against other sources – an impression confirmed in carrying out the case studies. Data on human resources for evaluation were also absent for a number of countries.

Despite data deficiencies, certain insights can still be gained. Where figures for *evaluation staff* at national level were provided, they ranged from one to five (most frequent responses being two and three), with a few exceptions where this number was significantly higher (e.g., in case of Italy, the Evaluation Unit at the Ministry of Economy was reported to have 15 staff members). Information on the number of staff at programme level was more difficult to establish because responses were provided by representatives of just some OPs for some countries. But where numbers were provided, staff numbers also ranged between one and five.

On the whole, survey and case study findings confirm that financial resources are not seen as a constraint on the evaluation of Structural and Cohesion Fund programmes. A clear majority of respondents on this issue thought that financial resources allocated for the evaluation function were adequate. In contrast, only 38 percent of respondents indicated that human resources were adequate.

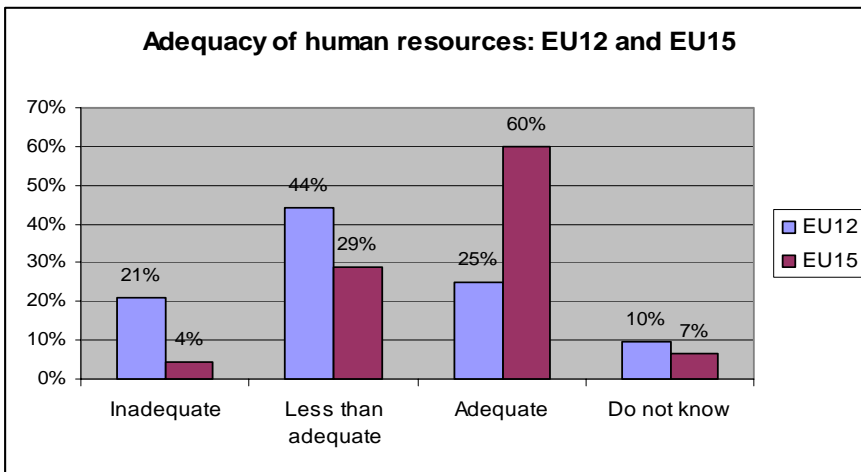


n=118

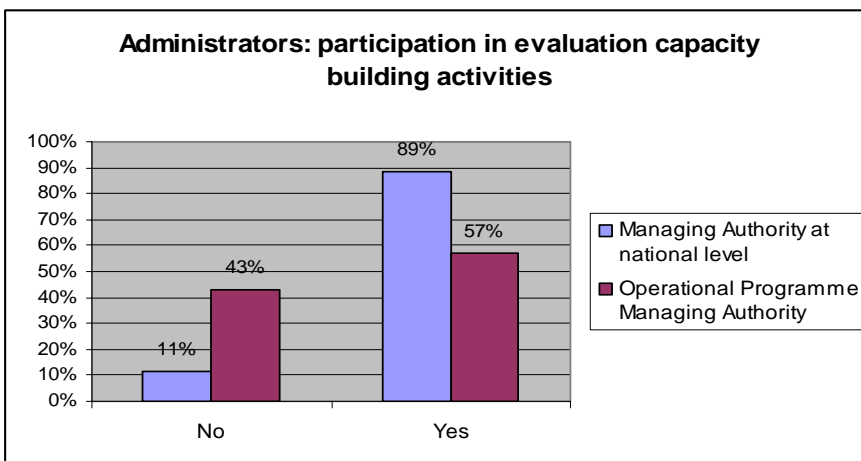


n=118

There are notable differences between responses from the EU12 and EU15 on the adequacy of human resources. While 60 percent of respondents from the EU15 think that human resources for evaluation are adequate, only 25 percent respondents from the EU12 have the same opinion. These findings suggest that the issue of human resources for evaluation deserves more attention, especially in the EU12.

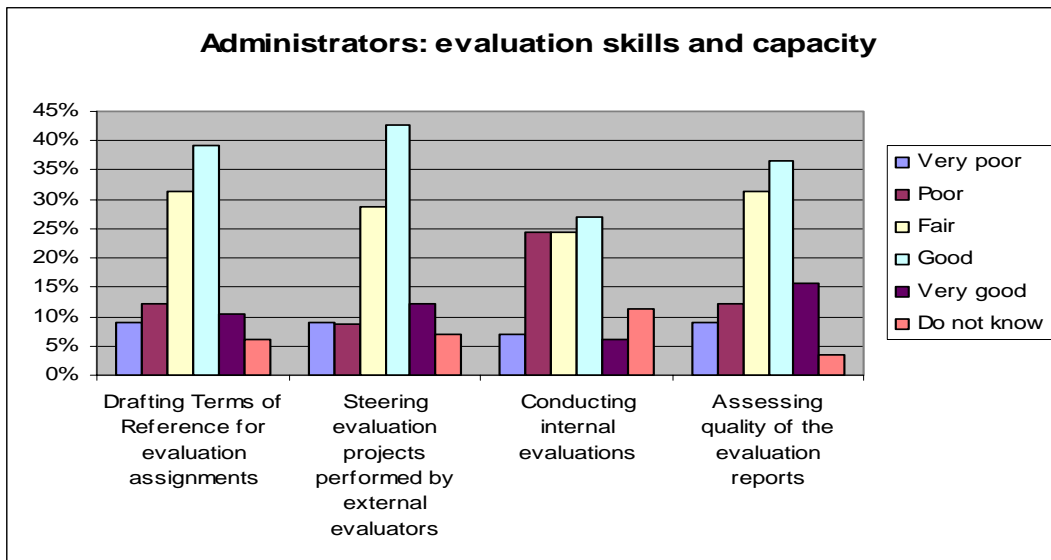


As regards *opportunities to strengthen evaluation capacity* within administrations through specialised training and education, responses received were too contradictory to allow for any clear-cut conclusions on their adequacy or otherwise. The most frequently mentioned evaluation capacity building events carried out in the period from 2004 to 2007 were workshops/seminars and training activities, but other measures such as evaluation guides and awareness raising activities were also cited. 70 percent of administrators who were aware of evaluation capacity building events had participated in them. In most cases these were workshops/seminars and training. A higher proportion of respondents from the managing authorities at national level have participated in capacity building events compared with those at the programme managing authorities.



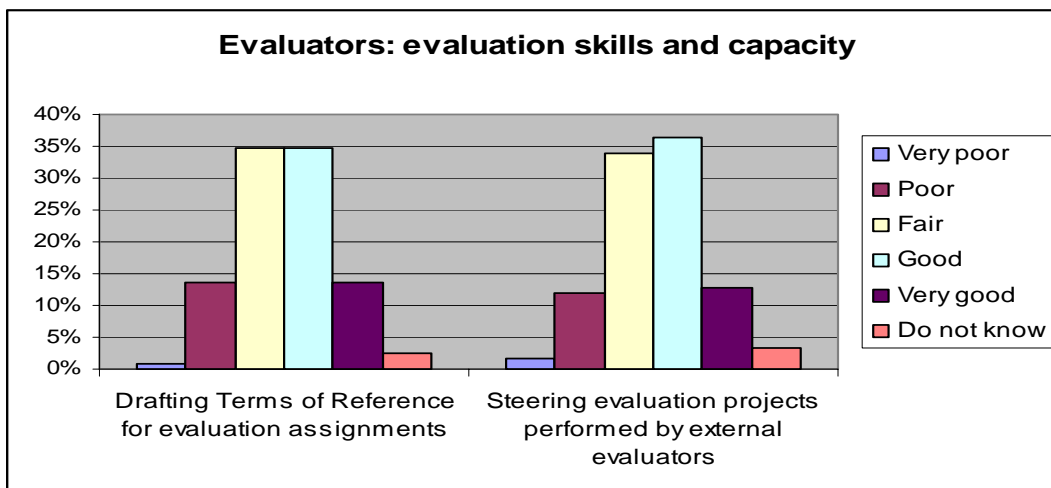
n=117

Asked to assess *skills and capacity*, most respondents from Structural and Cohesion Fund administrations indicated that skills in *drafting terms of reference* for evaluation assignments are either good or fair. Skills in *steering evaluation projects* performed by external evaluators are also seen as good or fair. A similar assessment was provided regarding skills for *assessing the quality* of evaluation reports. However, the assessment as regards skills in *conducting internal evaluations* was less positive. This suggests that officials believe that they are generally better skilled at managing the evaluation process than doing evaluations themselves. As this is a self-assessment the feedback has to be considered cautiously. The case studies carried out in the context of this project indicate that EU12 authorities, especially at programme level, still face certain difficulties in drafting high-quality terms of reference for individual evaluation assignments and ensuring quality control of external evaluation services.



n=115

To put the self-assessment provided by administrations in context, a similar question about evaluation capacity and skills in the Structural and Cohesion Fund administration was put to evaluation market participants. Results reveal that the assessment of skills in both drafting terms of reference and steering evaluation projects provided by market participants is quite similar to that of administrators and rather positive.



n=118

Efficiency of Evaluation Management

In order to assess the efficiency of evaluation management, issues relating to the planning, procurement, steering and quality control of evaluation projects were addressed.

Data on *evaluation plans* was not obtained for every country. In spite of that, some general observations regarding evaluation planning can be made. In the 2000-2006 period a national evaluation plan was drafted in a number of Member States. For the EU15 it was more common to have an evaluation plan at programme level. For the 2007-2013 programming period evaluation plans will be more commonplace - in a majority of Member States at either national or programme level. In many cases evaluation plans will be prepared at both levels. It would appear that the recommendations of the European Commission concerning planning of evaluation activity in the new period played a role in this respect.

Table 11. Evaluation plans

EU Member State	2000-2006 programming period		2007-2013 programming period	
	<i>Evaluation plan at national level</i>	<i>Evaluation plan at programme level³³</i>	<i>Evaluation plan at national level</i>	<i>Evaluation plan at programme level³⁴</i>
Austria	n.a.	no	in preparation	in preparation
Belgium	no	no	no	yes
Bulgaria	-	-	in preparation	yes
Cyprus	yes	n.a.	n.a.	n.a.
Czech Republic	n.a.	n.a.	yes / in preparation	yes / in preparation
Denmark	no	yes	In preparation	In preparation
Estonia	yes	n.a.	yes / in preparation	yes / in preparation
Finland	no	yes	in preparation	in preparation
France	n.a.	yes	in preparation	in preparation
Germany		yes	yes / in preparation	yes / in preparation
Greece	n.a.	yes	yes	yes / in preparation
Hungary	no	yes	yes / in preparation	yes / in preparation
Ireland	yes	n.a.	n.a.	n.a.
Italy	yes	no	in preparation	no
Latvia	n.a.	n.a.	n.a.	n.a.
Lithuania	yes	n.a.	yes / in preparation	no
Luxembourg	no	yes	n.a.	in preparation
Malta	yes	n.a.	in preparation	no
Netherlands	n.a.	yes	n.a.	yes / in preparation
Poland	yes	yes	in preparation	yes / in preparation
Portugal	n.a.	yes	in preparation	in preparation
Romania	-	-	yes / in preparation	yes / in preparation
Slovakia	n.a.	n.a.	in preparation	n.a.
Slovenia	yes	n.a.	yes	yes
Spain	n.a.	yes	yes / in preparation	yes / in preparation
Sweden	n.a.	yes	no	no
United Kingdom	no	yes	no	in preparation

n.a. - not available

When asked about time needed to carry out the *procurement procedure* (from the point at which agreement is reached on terms of reference for an evaluation to the appointment of an evaluator), only half of all respondents answered with the duration ranging from 3 to 60 weeks. The case studies indicate that the length of the procurement process depends on the scope and budget of the evaluation assignment.

As already mentioned, evaluation *steering groups* or committees are one of the most common tools to steer evaluation activity in the EU Member States. Steering groups *at programme level* are set up in a majority of the EU Member States. However, because of contradictory responses and non-responses, data on the existence of evaluation steering groups *at national level* are not available for almost half of the EU Member States or for an even larger share in respect of *individual projects*. The Member State case studies indicate that steering groups differ in composition, i.e., in terms of whether they include external participants, such as representatives from voluntary sector, evaluation experts and the like.

³³ Indicated „Yes“, if at least one OP in the country has an evaluation plan.

³⁴ Indicated „Yes“, if at least one OP in the country has an evaluation plan.

Table 12. Evaluation steering groups

EU Member State	Evaluation steering group at national level	Evaluation steering group at programme level ³⁵	Evaluation steering group for individual projects
Austria	yes	no	no
Belgium	no	yes	n.a.
Bulgaria	n.a.	yes	n.a.
Cyprus	yes	no	yes
Czech Republic	n.a.	yes	n.a.
Denmark	no	no	no
Estonia	yes	n.a.	n.a.
Finland	no	yes	no
France	n.a.	yes	n.a.
Germany	yes	yes	n.a.
Greece	n.a.	yes	n.a.
Hungary	n.a.	yes	yes
Ireland	yes	yes	yes
Italy	no	n.a.	n.a.
Latvia	n.a.	n.a.	n.a.
Lithuania	yes	n.a.	yes
Luxembourg	n.a.	yes	no
Malta	yes	no	no
Netherlands	n.a.	yes	n.a.
Poland	yes	yes	n.a.
Portugal	n.a.	yes	yes
Romania	n.a.	yes	n.a.
Slovakia	n.a.	yes	no
Slovenia	yes	n.a.	n.a.
Spain	yes	yes	n.a.
Sweden	no	yes	n.a.
United Kingdom	n.a.	yes	n.a.

n.a. - not available

Finally, in order to assess the issue of *quality control*, survey and case studies explored the existence of approved quality standards for the assessment of evaluation reports. Survey responses received were contradictory or non-existent for more than a half of the EU Member States, which undermines the possibility of straightforward conclusions. The case studies demonstrate that some institutions develop quality standards on the basis of EVALSED quality assessment criteria (which are also referred to in the Commission's methodological papers³⁶), while others develop their own criteria or rely on research and evaluation expertise rather than pre-set formal criteria in judging the quality of evaluation reports.

Quality of Monitoring Systems

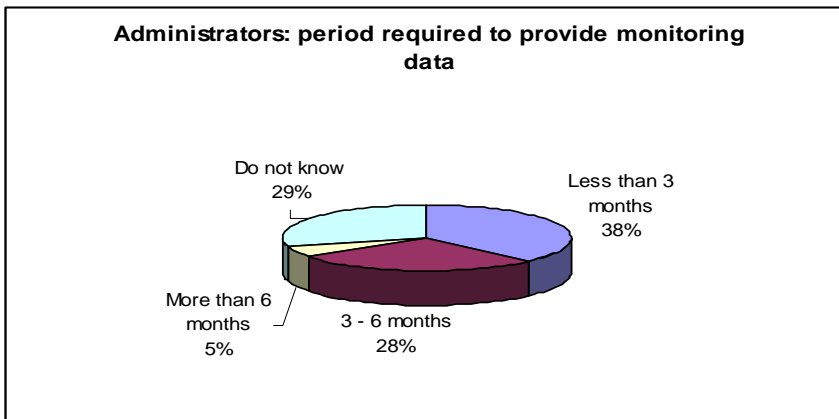
The survey and case studies attempted to explore the issue of *quality* of Structural and Cohesion Fund *monitoring systems* and timeliness of indicator data. There is a clear divergence in opinions

³⁵ Indicated „Yes“, if at least one OP in the country has an evaluation steering group set up.

³⁶ For example, Annex 2 of European Commission's Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5.

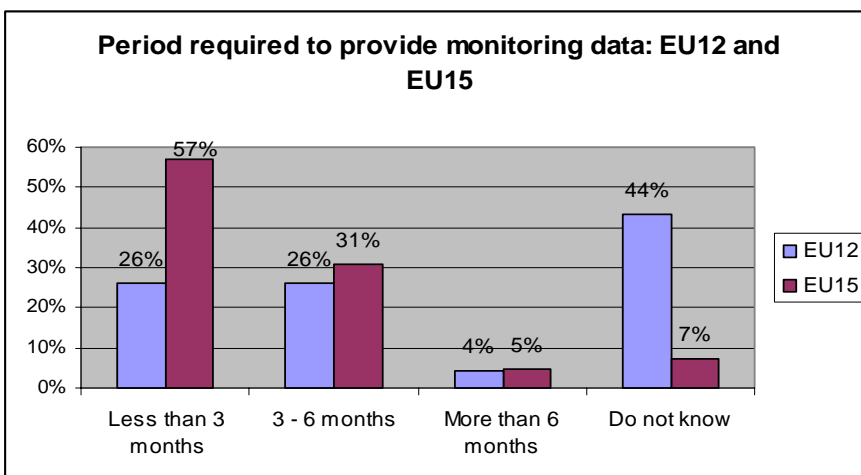
http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf

expressed by survey respondents regarding the period required for the monitoring system to provide data for the previous year. Nearly 40 percent of administrators who responded indicated that it took less than 3 months; almost one third said that it took between 3 to 6 months and a similar proportion indicated that they did not know the answer.

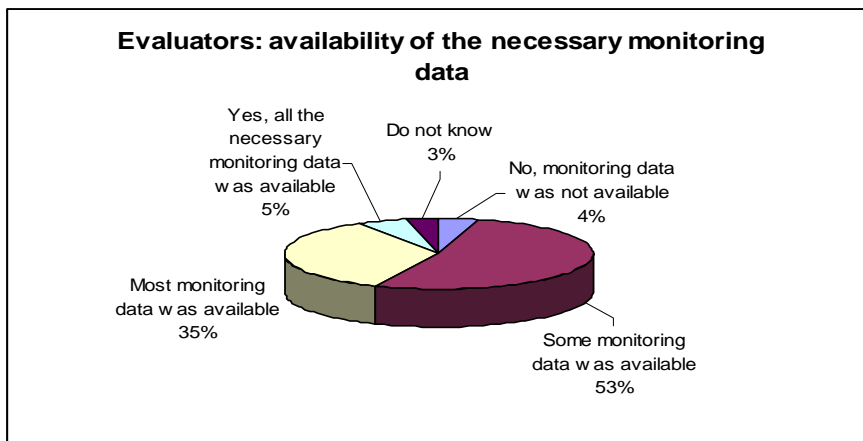


n=112

There is a noticeable variation on this issue between respondents from the EU15 and EU12. More than a half of respondents (57 percent) from the EU15 indicated that it took less than 3 months to obtain monitoring data for the previous year compared with only 26 percent of administrators from the EU12. Almost half of respondents from the EU12 did not know the answer to this question. In contrast this percentage was only 7 percent for the EU15.



Evaluation market participants were also asked to assess the availability of the necessary monitoring data for their most recent Structural and Cohesion Fund programme evaluation. More than a half of those who answered the question said that *some* monitoring data was available and more than one third of them said that *most* monitoring data was available.



n=118

Case studies of the selected Member States also indicate that in most cases there is still room for improvement in the monitoring systems – i.e. in terms of the definition of indicators, their timeliness and consistency across programmes and measures. Against this background, initiatives undertaken by some Member States to evaluate the functioning of monitoring systems and possibilities for their improvement can be considered as good practice examples.

Summary of findings concerning the demand side

Summing up analysis of the demand side, the following observations can be made.

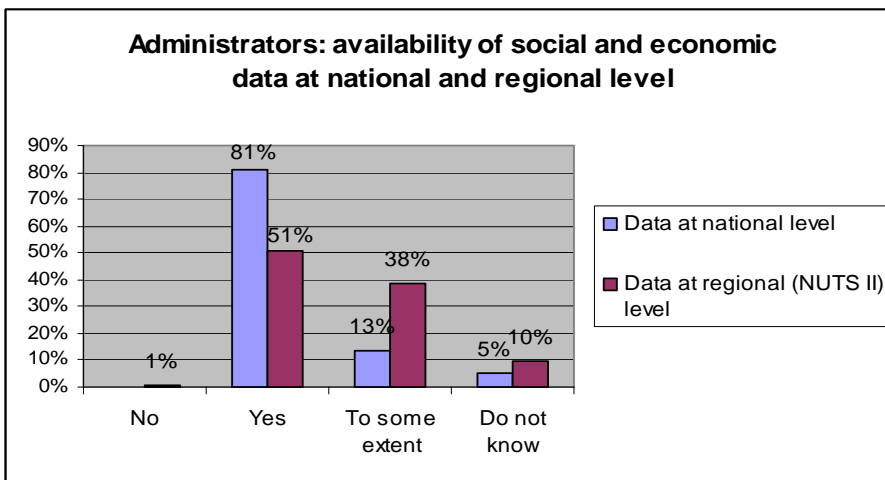
- The dominant practice in the Member States is to set up formal structures for the evaluation of Structural and Cohesion Funds and establish formal mechanisms for coordination of evaluation at national level. The most common instruments for this purpose are steering groups and evaluation units.
- Views of evaluation market participants with regard to definition of institutional responsibilities and procedures for Structural and Cohesion Fund evaluation function are largely positive. However, evaluators from the EU-12 are more critical in this respect.
- Evaluation planning is becoming a common practice in the 2007-2013 programming period; this is an encouraging finding and should help in managing evaluation demand in the 2007-2013 period.
- Both the survey and case studies indicate that administrators are not well-informed as regards the level of financial resources allocated to the evaluation function or that this information is not readily at hand. In spite of that, financial allocations for the evaluation of Structural and Cohesion Fund programmes are seen as generally adequate in contrast to human resources. As a rule, Structural and Cohesion Fund programmes have 1 to 5 persons responsible for the evaluation function - a level which, according to many administrators, is less than adequate, especially for carrying out internal evaluations. Significant differences exist between EU15 and EU12 with the adequacy of human resources perceived as as a much bigger problem in the EU12.
- Survey evidence suggests that quality of monitoring systems varies between Member States. A majority of Structural Fund administrators from EU15 indicate that monitoring indicators for the previous year become available in less than three months. However, a large share of respondents from the EU12 could not answer this question. This is an indication of the short experience in programme monitoring and should also be taken into account when considering the issue of linkages between monitoring and evaluation.

5.1.3. The Supply Side Dimension

Having discussed the demand-side criteria and indicators, this section focuses on the supply-side dimension of evaluation systems, which is assessed on the basis of two criteria: quality of socio-economic data resources and the availability and quality of evaluation supply on the domestic market.

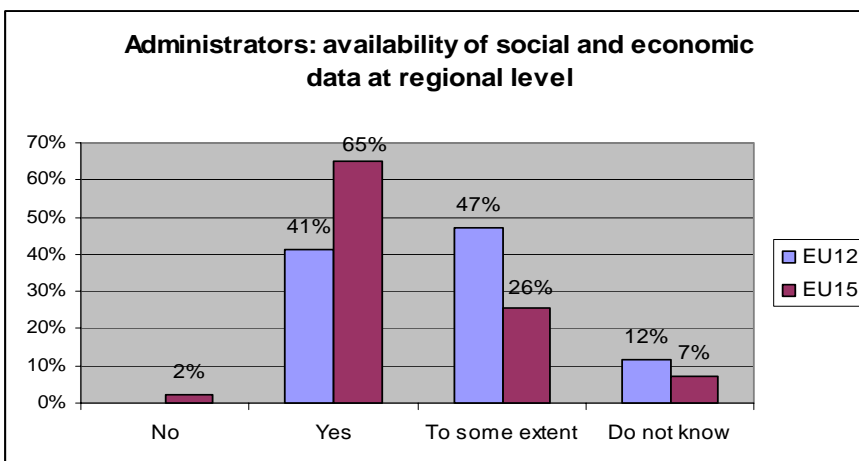
Quality of Socio-Economic Data

As shown in the graphs below, a large majority of respondent administrators were of the view that key statistical economic and social *data on the national level* is available to evaluators (although a response of the form “data available to some extent” featured to a greater extent in the EU12). However, *data on regional (NUTS II) level* was not regarded as being as widely available. These findings are in line with case study results.

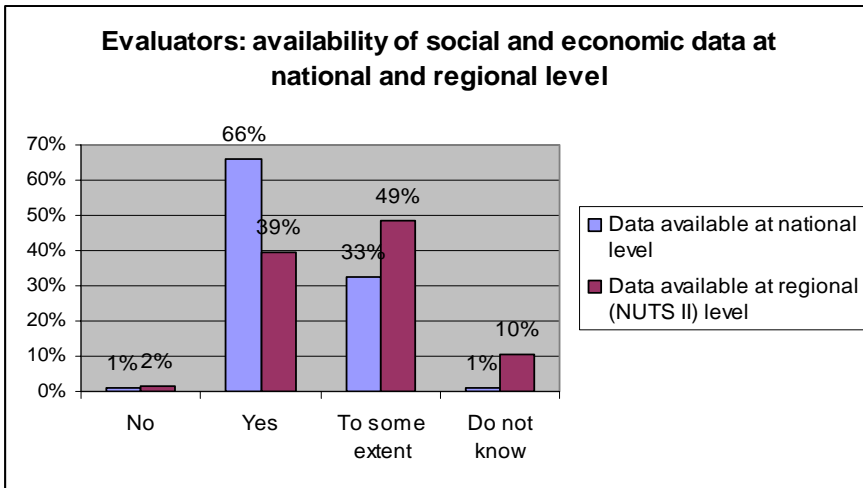


n=112

Survey results reveal that there are differences in the responses as between officials in the EU15 and EU12. While 65 percent of respondents from the EU15 indicate that data on regional level is available, this was the opinion of 41 percent of EU12 administrators.



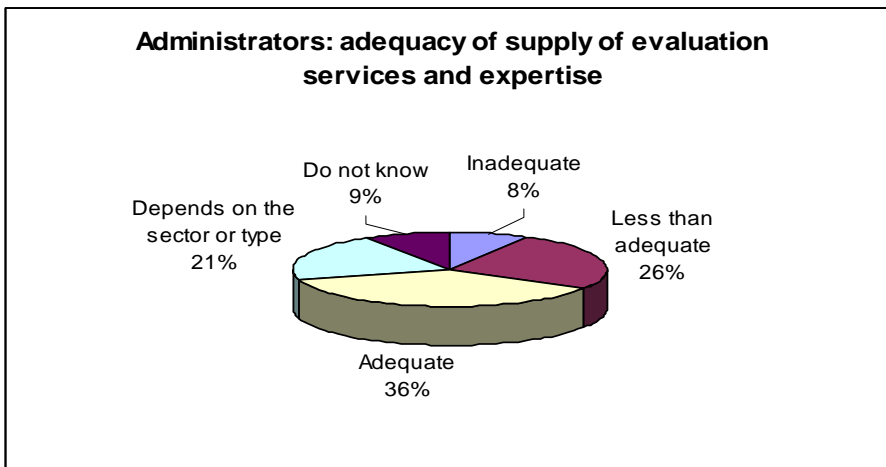
These views were shared by evaluation market participants with a majority being of the view that key official statistical data on the *national level* was available and one third that data was available to some extent. Almost half of evaluators thought that only some *regional level* data was available.



n=117

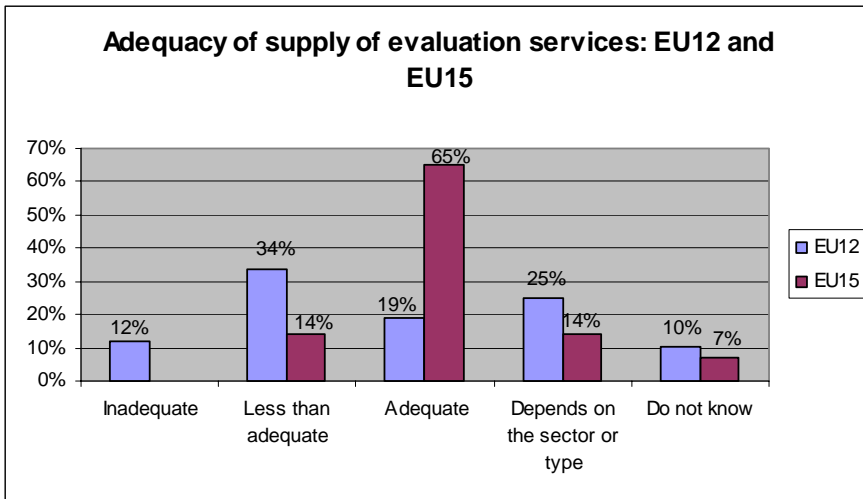
Availability and Quality of Evaluation Expertise

Asked about the *adequacy of evaluation services and expertise* in the national market, only 37 percent of responding administrators indicated that the supply of such services and expertise was adequate.



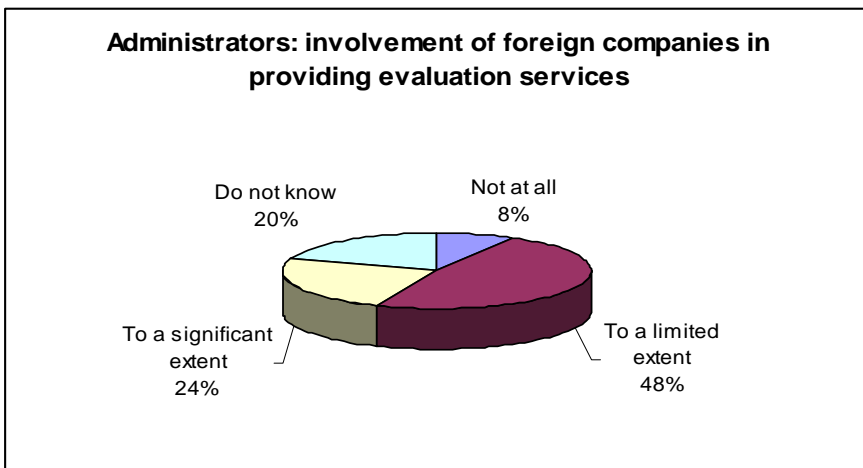
n=112

There is a significant difference between the responses from EU15 and EU12. Only one fifth of administrators from the EU12 are of the opinion that the supply of evaluation services is adequate compared with 65 percent for the EU15. Views to the effect that the supply of evaluation services was inadequate were received from the EU12 only. This confirms that evaluation markets in many of the EU12 countries are still under-developed, a view that the case studies would broadly confirm.

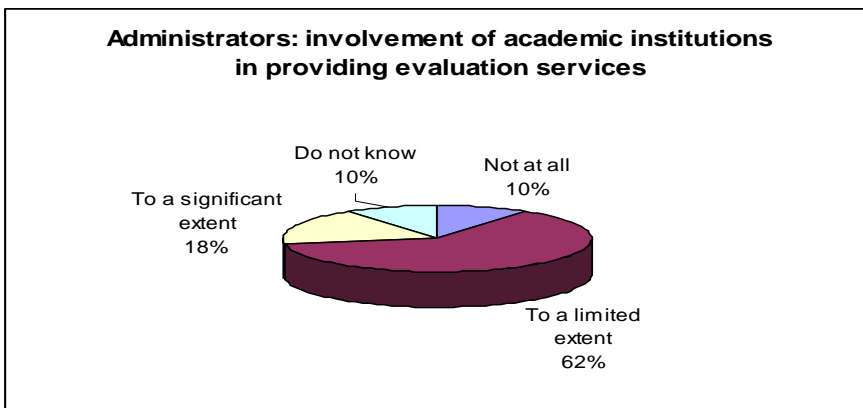


Half of the respondents indicated that *foreign companies* were involved in providing evaluation services to a limited extent. While responses for individual Member States have to be considered cautiously because of sample size, a few countries stand out as regards a reported significant involvement of foreign companies – i.e. the two newest EU Member States (Bulgaria and Romania), which started developing Structural and Cohesion Fund evaluation capacity only recently.

In overall terms, the involvement of *academic institutions* was judged to be of a limited extent according to 62 per cent of respondents.



n=112



n=112

In order to explore the issues of possible “*internal supply*” of evaluation within Structural and Cohesion Fund administrations, administrators were asked about the existence of units carrying out

internal evaluations of “their” programme. Survey results do not allow the existence/non-existence of such units at national level to be established in four EU Member States, because responses from managing authorities at national level were not received. Survey results concerning the remainder suggest that internal evaluations at national level are carried out in six countries. The number of internal evaluations conducted by those institutions was not provided in all cases, but where indicated, ranged from 1 to 4. The Portuguese *Observatorio do QCA III* stands out in this regard as the number of internal evaluations reported is significantly higher – i.e. 37 evaluations. While analysis of responses concerning internal evaluations at programme level is once again complicated due to patchy survey coverage of programme level institutions, internal evaluations at programme level are carried out at least in nine EU Member States. The number of such evaluations usually varies from 1 to 4.

Table 13. Internal evaluation within Structural and Cohesion Fund administrations

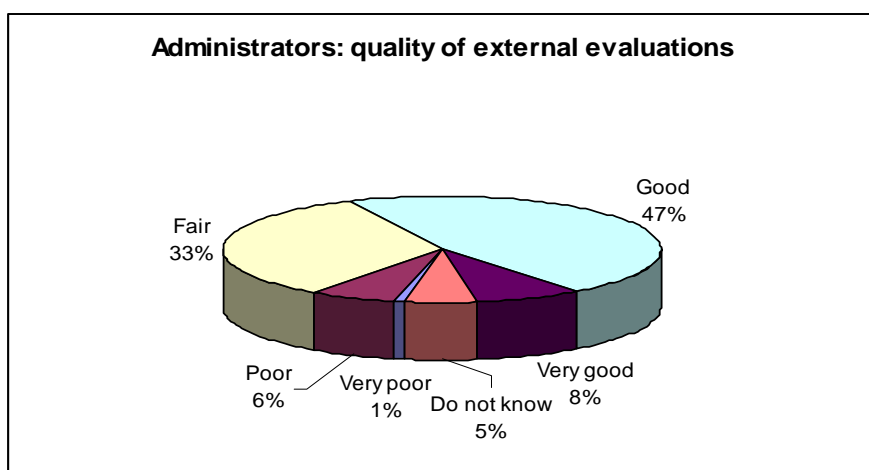
EU Member State	Internal evaluation unit at national level		Internal evaluations at programme level institutions ³⁷
	Institution	Number of internal evaluations 2004-2006	
Austria	no	-	n.a.
Belgium	no	-	n.a.
Bulgaria	no	-	yes
Cyprus	no	-	n.a.
Czech Republic	Evaluation and analysis unit at the Ministry for Regional Development (MA at national level)	1	yes
Denmark	no	-	n.a.
Estonia	no	-	n.a.
Finland	no	-	n.a.
France	no	-	yes
Germany	n.a.	n.a.	yes
Greece	no	-	n.a.
Hungary	Evaluation and Methodology Unit at the National Development Agency, which is MA at national level	n.a.	yes
Ireland	n.a.	n.a.	n.a.
Italy	Evaluation Unit of the Department for Development Policies, Ministry for Economic Development, which is MA at national level	4	n.a.
Latvia	no	-	yes
Lithuania	no	-	
Luxembourg	<ul style="list-style-type: none"> ▪ Yes in Ministry of Economy and Foreign Trade ▪ No in Ministry of Labour and Employment 	<ul style="list-style-type: none"> ▪ 2 ▪ -- 	n.a.
Malta	no	-	n.a.
Netherlands	no	-	n.a.
Poland	no	-	n.a.
Portugal	Observatorio do QCA III under the Ministry for Environment,	37	n.a.

³⁷ Indicated „Yes“, if at least one institution at OP level in the country carries out internal evaluations.

EU Member State	Internal evaluation unit at national level		Internal evaluations at programme level institutions ³⁷
	Institution	Number of internal evaluations 2004-2006	
	Territorial Planning and Regional Development		
Romania	no	-	yes
Slovakia	no	-	n.a.
Slovenia	Government Office for Local Self-Government and Regional Policy (MA at national)	n.a.	n.a.
Spain	n.a.	n.a.	yes
Sweden	n.a.	n.a.	n.a.
United Kingdom	n.a.	n.a.	yes

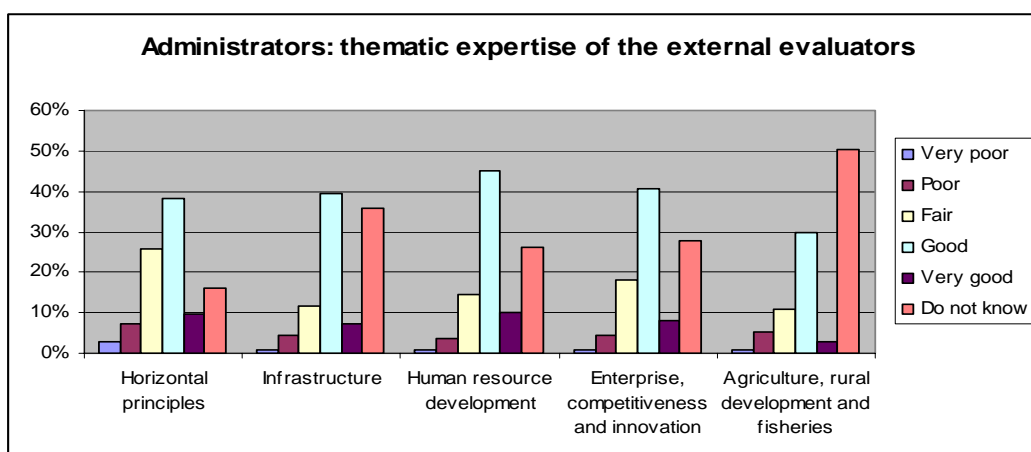
n.a. - not available

Administrators were asked to assess the *quality of evaluation reports* carried out by the external evaluators in recent years. As it may be seen from the graph below, their assessment is rather positive. A “very good” rating featured more frequently in the EU15. An assessment of reports as “poor” was noted for three Member States and “very poor” in the case of just one Member State.



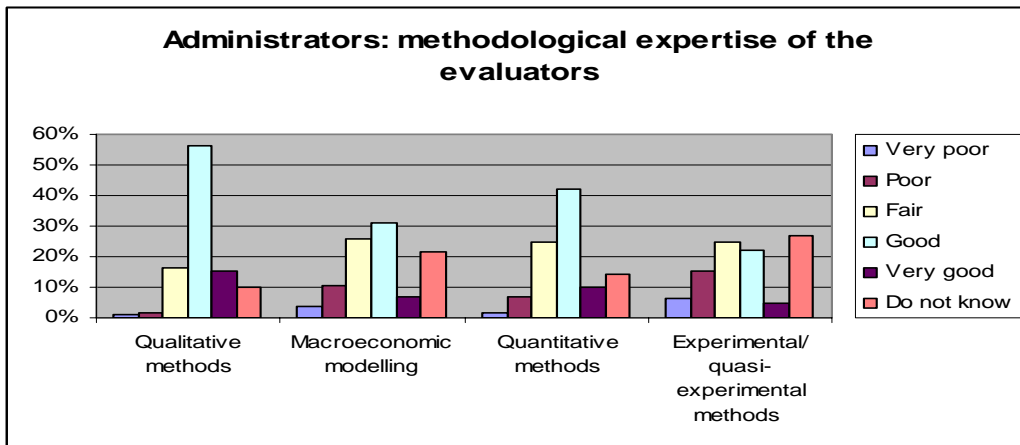
n=112

In terms of *thematic expertise*, administrators gave a positive appraisal to external evaluators providing services in the areas of human resource development and enterprise, competitiveness and innovation. Expertise in horizontal themes was judged to be slightly worse. Expertise in the field of agricultural, rural development and fisheries was judged to be the lowest, although half of administrators could not provide their opinion on the expertise in this area.



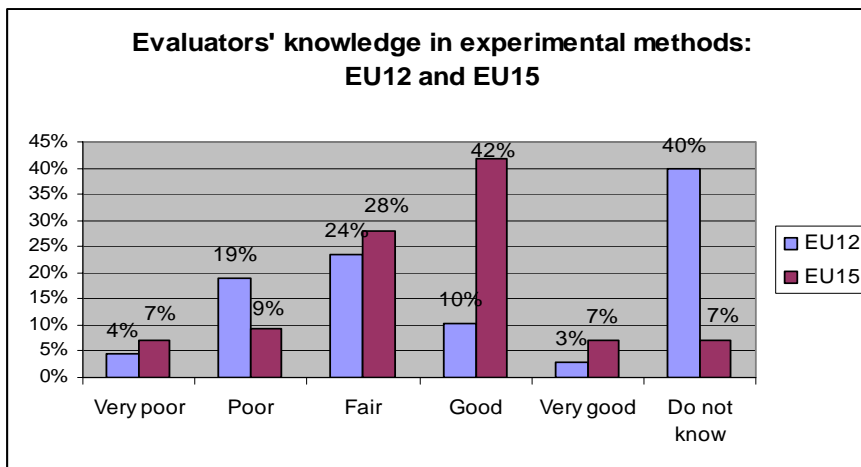
n=111

As regards *methodological expertise*, the use of qualitative methods was judged to be the best. Expertise in quantitative methods was also been judged to be quite good. Expertise in macroeconomic modeling was considered to be lower and even more so in experimental methods, although about one quarter of respondents could not provide an opinion on this aspect.

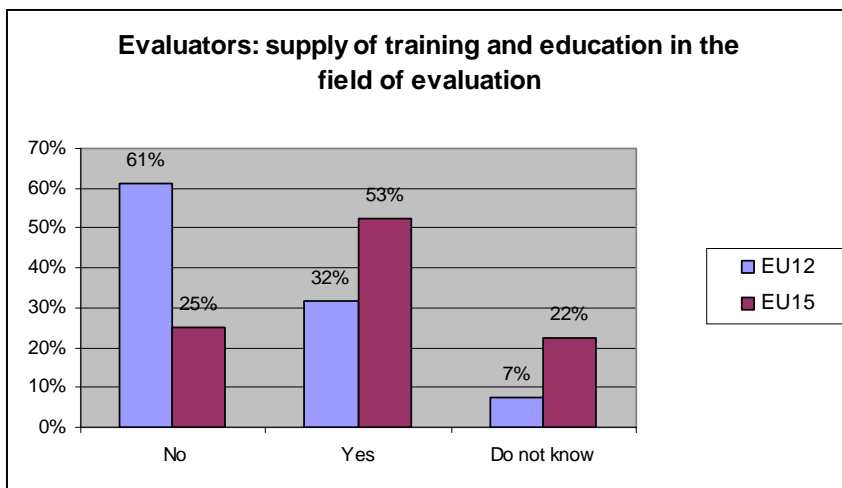


n=112

Administrators from the EU15 are more positive about evaluators' expertise in the area of experimental methods than their counterparts in EU12: 42 percent of respondents from the EU15 assessed this expertise as good, while only 10 percent of administrators from the EU12 had the same opinion. However, almost 40 percent of respondents from the EU12 did not offer an opinion on this question.



Evaluation market participants were asked about the *supply of training and education* in the field of evaluation in their country. The results indicate that there are differences between training opportunities in the EU15 and EU12. A majority of respondents from the EU12 stated that that there was no training and education supply in evaluation, whereas a majority of evaluators from the EU15 indicated that there was.



n=117

As the case studies illustrate, a range of well-received evaluation capacity building measures supported by the pre-accession financing (Phare) were carried out in the EU12 to create the initial evaluation capacity both on the demand and supply sides. Case studies also suggest that evaluation societies/associations tend to play an important role in organizing the supply of evaluation training and other capacity building events for evaluators.

Summary of findings concerning the supply side

Summing up analysis of the supply side, the following observations can be made.

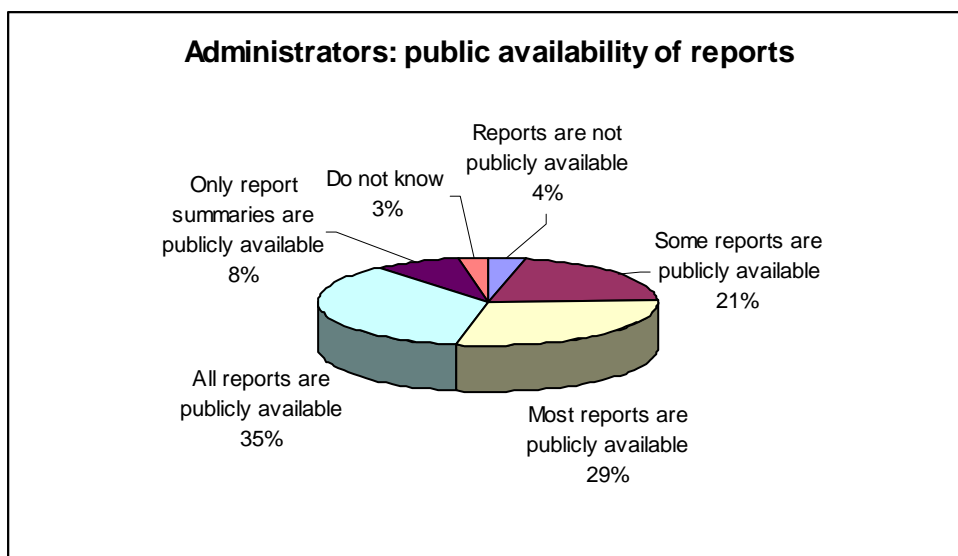
- The most important feature regarding the quality of socio-economic data for evaluations is that data on the national level is more readily available than on the regional level. Both administrators and market participants tend to agree on this.
- Internal evaluation capacity within Structural and Cohesion Fund administrations is limited. 'In-house' evaluations at national level are carried out in just a few Member States, with programme level internal evaluations slightly more common. While evaluation units which carry out internal evaluations do exist in some countries, the predominant approach across the Member States is to commission evaluations from external sources.
- Survey and case study results support the widely-held opinion that evaluation markets in many of the EU12 countries are still under-developed and that the supply of evaluation services is still inadequate; for example only one fifth of administrators from the EU12 are of the view that the supply of evaluation services is adequate.
- With few exceptions, the results indicate that foreign companies, universities and research institutes are involved to a limited extent in the provision of evaluation services. The survey suggests that the quality of external evaluation is considered to be quite good, in particular in the fields of human development, enterprise, competitiveness and innovation. The use of both qualitative and quantitative methods is judged to be quite good, while expertise in the use of experimental methods and macroeconomic modeling is felt to be lower, in particular in the EU12.
- Survey results suggest that training and education opportunities in the evaluation field are more limited in the EU12.

5.1.4. The Dissemination and Utilisation Dimension

Dissemination and utilisation aspects were assessed on the basis of the public availability of evaluation reports, the institutions and procedures for dealing with evaluation findings and the influence of the most recent ex-ante evaluations. Both administrators and evaluation market participants were surveyed on these issues.

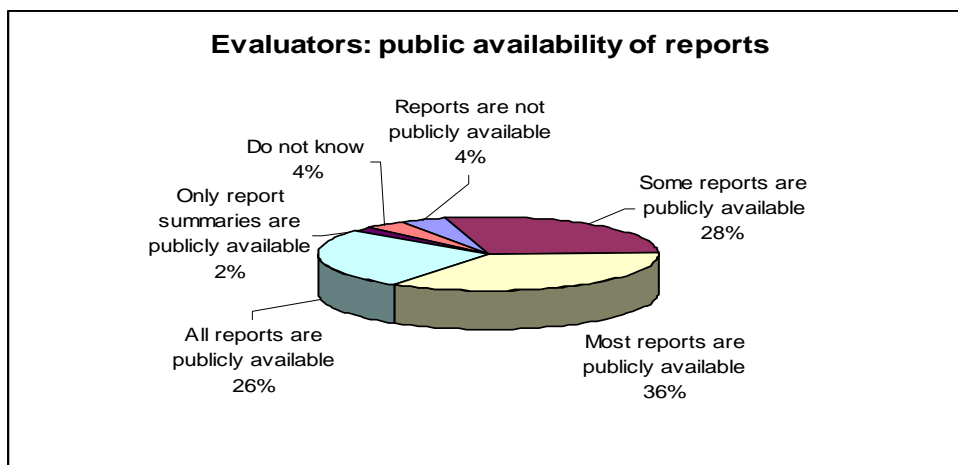
Dissemination

Nearly all surveyed administrators indicated that Structural and Cohesion Fund evaluation reports were publicly available. However, responses diverged considerably about the extent of dissemination – i.e. whether all, some or most reports were publicly available. This is illustrated in the graph below.



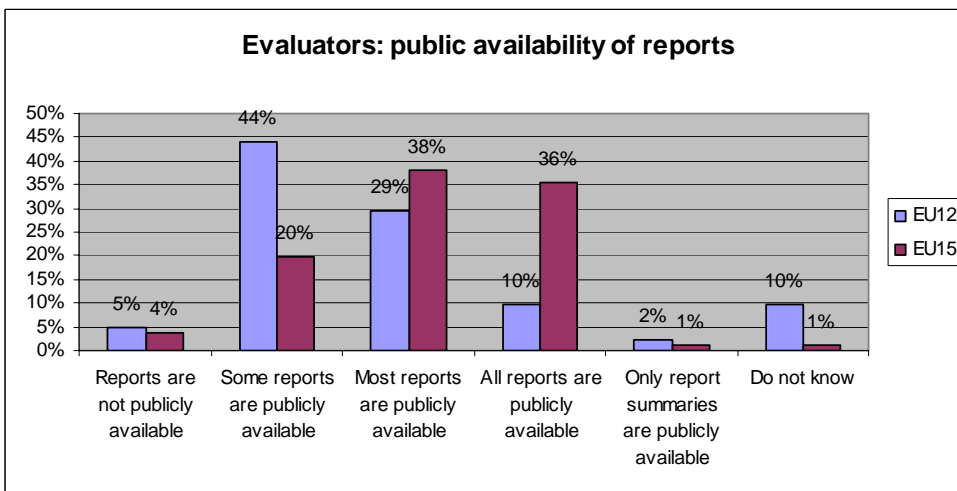
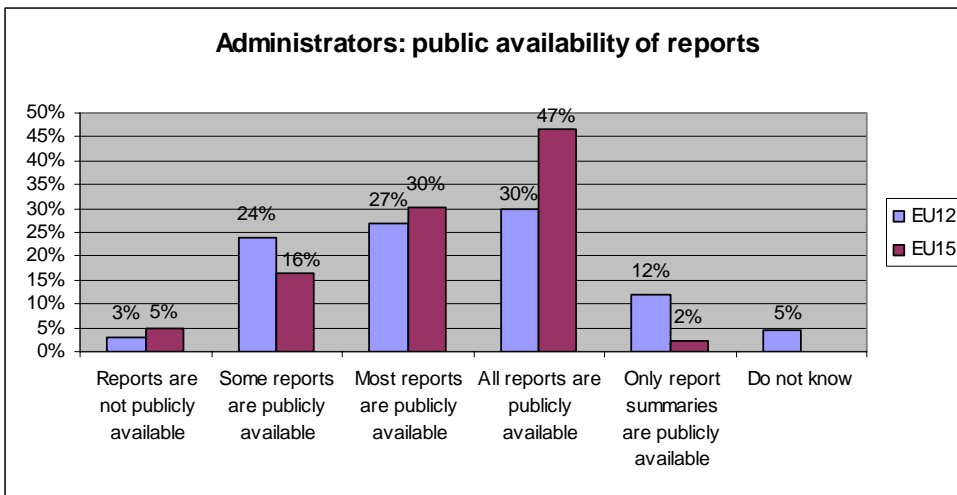
n=111

Evaluation market participants expressed quite similar views regarding the public availability of evaluation reports.



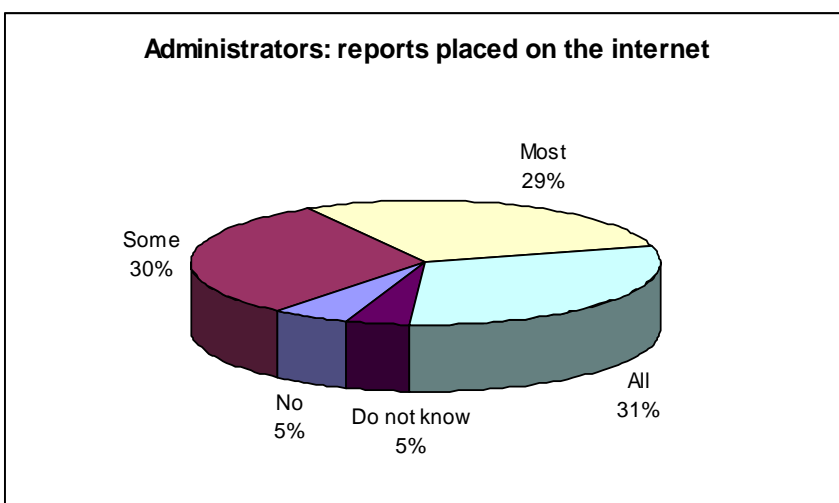
n=117

It can be observed that a higher proportion of both administrators and evaluators from the EU15 state that *all* or *most* evaluation reports are publicly available.



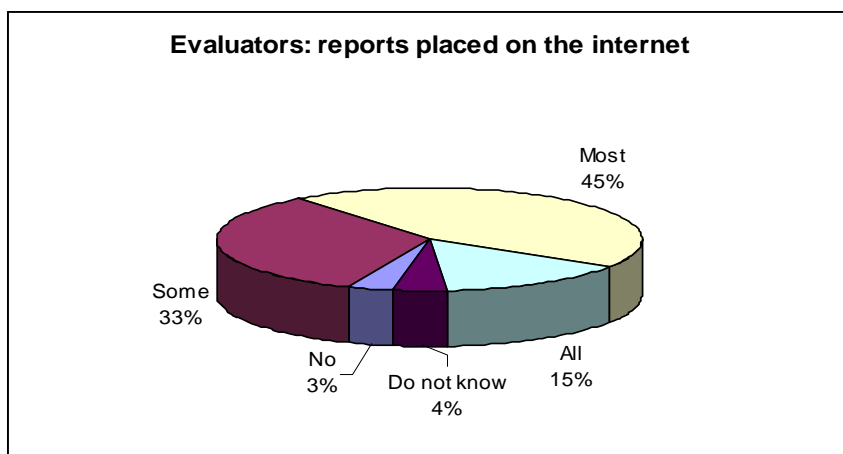
Case study evidence from one of the selected countries (Romania) illustrates that Member States may decide to set formal minimum requirements regulating the dissemination of evaluation reports.

With regard to online access to Structural and Cohesion Fund evaluation reports, a mixed picture emerges. Almost one third of administrators said that *all* reports were placed on the internet; another third indicated that *some* reports were placed on the internet and slightly less than one third responded that *most* reports are placed on the internet.



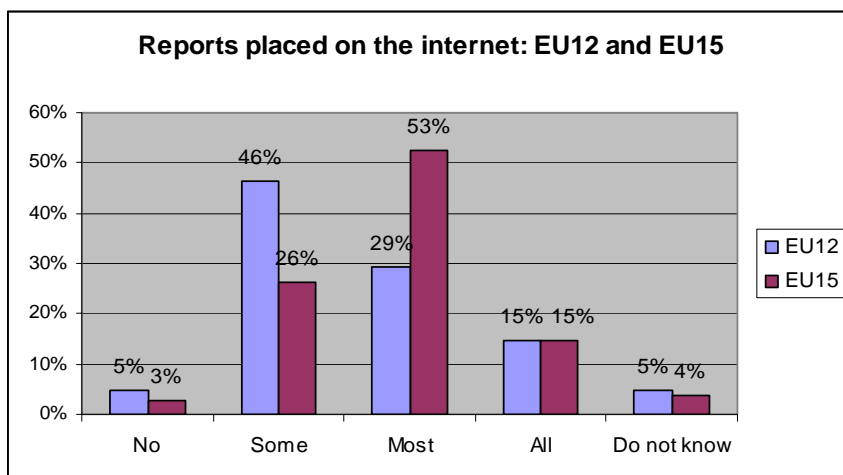
n=111

In contrast only 15 percent of market participants who responded indicated that *all* reports were available on the internet.



n=117

A majority of evaluators from the EU15 thought that most reports are placed on the internet, whereas evaluators from EU12 most often chose answer “some reports” available on the internet.



In order to assess the *utilisation* criterion, the existence of institutions with responsibility for taking forward the findings of completed evaluation reports as well as procedures for following-up of agreed recommendations were considered. Table 14 below specifies responsible institutions per country as indicated in survey responses. While the information received is not altogether clear or its interpretation straightforward as regards the roles and division of responsibilities between the institutions mentioned, certain observations can be made. It is noticeable that typically the institutions with responsibility for considering or taking forward evaluation findings include monitoring committees, managing authorities, evaluation units and evaluation steering groups. In some cases the addressees of evaluation recommendations are responsible for dealing with the recommendations. It can also be observed that it is not uncommon for programme level institutions to have their own specific arrangements for considering or taking forward evaluation findings.

Table 14. Institutions and procedures for addressing evaluation results

EU Member State	Institution responsible for considering and taking forward evaluation findings		Procedures for following-up the implementation of agreed recommendations	
	<i>as indicated by MA at national level</i>	<i>as indicated by programme level institutions</i>	<i>at national level</i>	<i>At programme level³⁸</i>
Austria	n.a.	n.a.	n.a.	no
Belgium	n.a.	n.a.	n.a.	
Bulgaria	Monitoring Committees	<ul style="list-style-type: none"> ▪ MA ▪ Monitoring Committee 	yes	yes
Cyprus	Evaluation Steering Committee	n.a.	no	n.a.
Czech Republic	Evaluation and analysis unit at the Department of European Funds of the Ministry for Regional Development	<ul style="list-style-type: none"> ▪ Monitoring Committee ▪ NSRF and OP level evaluation units ▪ Working group for evaluation of CSF 	no	n.a.
Denmark	no	n.a.	no	n.a.
Estonia	Evaluation Steering Group decides upon channelling evaluation recommendations and co-ordinates their follow-up.	n.a.	yes	yes
Finland	n.a.	Monitoring Committee	n.a.	yes
France	Monitoring Committee	<ul style="list-style-type: none"> ▪ A follow-up committee ▪ Each regional evaluation unit and the European Affairs Department of The Ministry for the Interior, Overseas and the Local Authorities ▪ Informal organisation 	no	yes
Germany	n.a.	<ul style="list-style-type: none"> ▪ MA ▪ Steering Group and Monitoring Committee together with MA ▪ KBS Public Contact Agency (Berliner Senatsverwaltung für Wirtschaft, Technologie und Frauen – OP MA) 	n.a.	yes
Greece	OP MAs	<ul style="list-style-type: none"> ▪ OP MA ▪ OP Monitoring Committee 	yes	yes

³⁸ Indicated „Yes“, if at least one institution at OP level in the country carries out internal evaluations.

EU Member State	Institution responsible for considering and taking forward evaluation findings		Procedures for following-up the implementation of agreed recommendations	
	<i>as indicated by MA at national level</i>	<i>as indicated by programme level institutions</i>	<i>at national level</i>	<i>At programme level³⁸</i>
Hungary	<ul style="list-style-type: none"> ▪ If a recommendation has an addressee institution, this is its task to deal with recommendations and disseminate them. ▪ New arrangements are foreseen for the 2007-2013 period 	Internal Control Unit	yes	yes
Ireland	n.a.	OP MA	n.a.	yes
Italy	no	n.a.	no	n.a.
Latvia	<ul style="list-style-type: none"> ▪ Monitoring Committee ▪ The Strategic Planning division of the Ministry of Finance ▪ Responsible persons of the Project Appraisal Department of Ministry of Finance 	<ul style="list-style-type: none"> ▪ Evaluation Committee (Ministry of Finance) ▪ Monitoring Committee ▪ Evaluation Steering Group is responsible for monitoring the process of implementation the recommendations of evaluation reports 	The Strategic Planning division of the Ministry of Finance prepares recommendation implementation plan and sends it out for update at least twice a year before Monitoring Committee meeting takes place.	yes
Lithuania	<ul style="list-style-type: none"> ▪ Institutions which initiated evaluations ▪ Evaluation division of the Ministry of Finance ▪ Recommendations are considered with responsible institutions 	Evaluation Division at the Ministry of Finance	no	yes
Luxembourg	<ul style="list-style-type: none"> ▪ Ministry of Labour and Employment: Steering committee ▪ Ministry of Economy and Foreign Trade: Follow-up committee 	n.a.	no	n.a.
Malta	The team coordinating the specific evaluation assignment.	n.a.	yes	n.a.
Netherlands	<ul style="list-style-type: none"> ▪ Ministry of Economic Affairs (Coordinator of the SF in Netherlands) and OP MAs. ▪ Ministry of Social Affairs and 	<ul style="list-style-type: none"> ▪ Monitoring committee ▪ MA ▪ OP MA 	<ul style="list-style-type: none"> ▪ Ministry of Economic Affairs - yes. ▪ Ministry of Social Affairs and 	n.a.

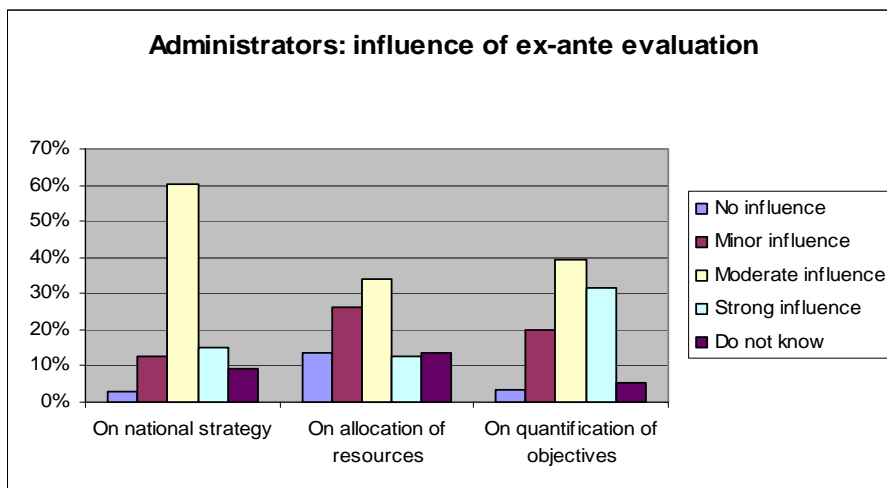
EU Member State	Institution responsible for considering and taking forward evaluation findings		Procedures for following-up the implementation of agreed recommendations	
	<i>as indicated by MA at national level</i>	<i>as indicated by programme level institutions</i>	<i>at national level</i>	<i>At programme level³⁸</i>
	Employment		Employment – no.	
Poland	Evaluation results are presented during the meetings (usually once in a month) where all stakeholders of the evaluation may discuss the results of the research.	<ul style="list-style-type: none"> ▪ Monitoring committee ▪ MA 	no	n.a.
Portugal	MAs	n.a.	yes	n.a.
Romania	<ul style="list-style-type: none"> ▪ Evaluation Central Unit ▪ MA at national level ▪ Evaluation Units ▪ OP MAs 	<ul style="list-style-type: none"> ▪ Monitoring Committee of the sectoral OPs ▪ OP Evaluation Unit ▪ Units within OP MAs 	A debriefing meeting is organised by the Evaluation Unit on the evaluation report findings and recommendations. All decision makers participate. After decision is taken on the implementation of recommendations (including means and schedule) the analysis of the status of their implementation becomes part of the monitoring process which is coordinated by the Monitoring Unit.	yes
Slovakia	no	<ul style="list-style-type: none"> ▪ Monitoring committee ▪ MA at national level 	no	yes
Slovenia	n.a.	n.a.	n.a.	n.a.
Spain	n.a.	Operational Programme Committee	n.a.	yes
Sweden	n.a.	<ul style="list-style-type: none"> ▪ Monitoring Committee 	n.a.	yes

EU Member State	Institution responsible for considering and taking forward evaluation findings		Procedures for following-up the implementation of agreed recommendations	
	<i>as indicated by MA at national level</i>	<i>as indicated by programme level institutions</i>	<i>at national level</i>	<i>At programme level³⁸</i>
		<ul style="list-style-type: none"> ▪ MA 		
United Kingdom	OP Monitoring Committee	<ul style="list-style-type: none"> ▪ ESF national monitoring committees for ESF evaluations ▪ Where an evaluation has been undertaken for a decision or makes appropriate recommendations the responsible unit is given the task to take appropriate action (South Yorkshire Objective 1 Programme Directorate, OP MA) ▪ Monitoring and Evaluation Working Group (EU Programme for Peace and Reconciliation and Interreg IIIA Programme – OP MA) 	yes	yes

n.a. - not available

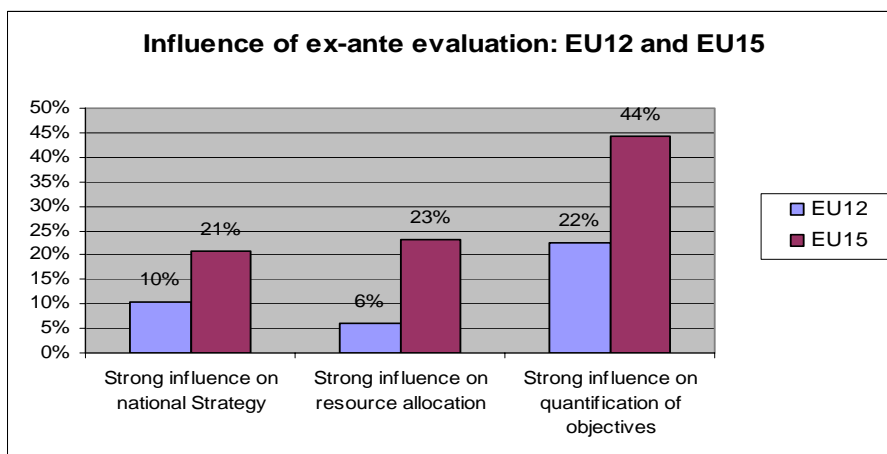
As regards *procedures for following-up or tracking of agreed evaluation recommendations*, clear-cut conclusions are difficult to make due to contradictory responses received from one third of Member States. Another third of Member States have certain procedures for following-up or tracking of agreed evaluation recommendations. Answers received are summarised in the table above. An interesting and unresolved question is the content of and arrangements for such procedures; however a survey is not the best method to pursue this detailed issue.

Asked about actual utilisation of recommendations of the most recent *ex ante evaluations*, more than 60 percent of responding administrators said that the evaluation had a moderate influence on the formulation of *national strategy* for Structural and Cohesion funding in 2007-2013. The influence of *ex ante evaluations* on the *allocation of resources* to different programs and priorities was judged to be less important; about one third of respondents believed the evaluations had a moderate influence, and 26 percent that they had a minor influence. The influence of *ex ante evaluation* on the *quantification of objectives and indicator selection* was judged to be comparatively important. These findings are fully in line with case studies' experience.



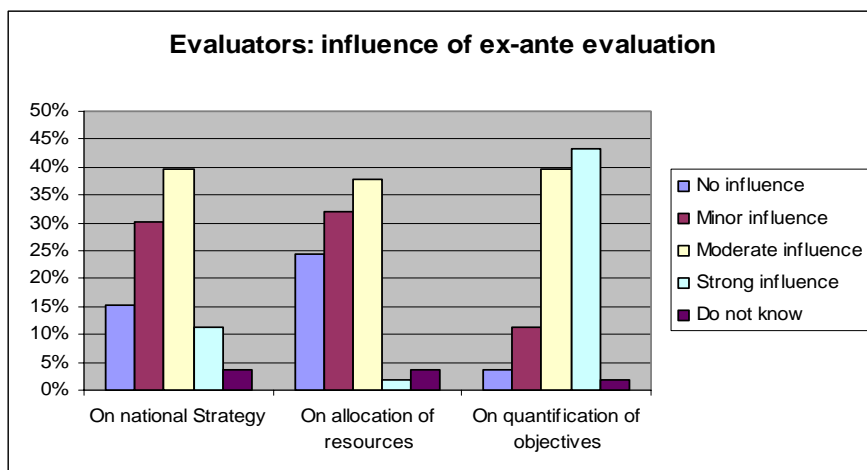
n=111

The survey feedback reveals interesting differences in the responses from the EU15 and EU12. A significantly greater proportion of respondents from the EU15 thought that *ex ante evaluation* had a strong impact on national Structural and Cohesion Fund investment strategy and quantification of objectives. Even greater differences can be observed with regard to resource allocation. This is illustrated in the graph below.



Evaluation market participants (those who have recently carried out *ex-ante evaluations*) were also asked to provide their opinion on the influence of evaluation recommendations. The results show

that evaluators, in particular from the EU12, are notably more sceptical in this regard than public authorities. Similar to the views of administrators, the greatest impact is felt to have been on the quantification of objectives.



n=53

Summary of findings concerning dissemination and utilisation

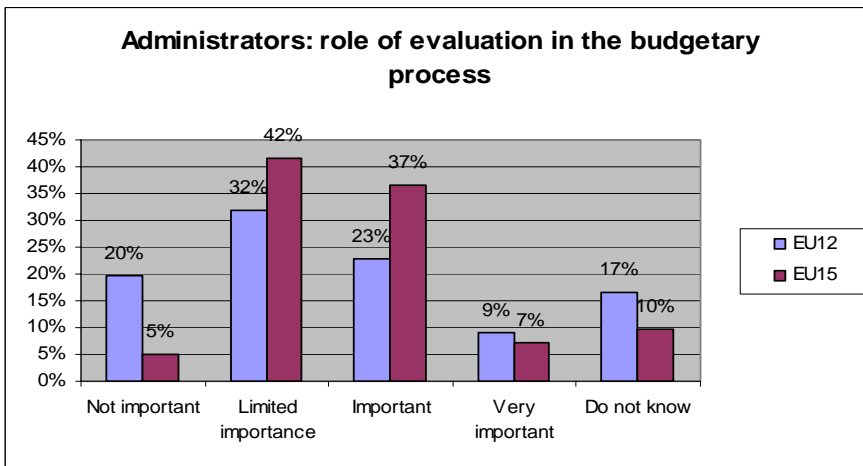
Summing up analysis of utilisation and dissemination dimension, the following observations can be made.

- Evaluation reports are quite widely available to the public, including via internet. Accessibility to reports (or at least awareness of them) is higher in the EU15.
- As a rule, there are certain arrangements in place to address evaluation findings. The most common institutional arrangements include monitoring committees, evaluation steering groups, evaluation units or managing authorities.
- However, assessment of the influence of the most recent ex-ante evaluation, which is closely integrated into the programming process, shows that utilisation of evaluation results is still limited. Comparing the degree of influence, the impact of ex-ante evaluations is judged to be the greatest on the quantification of objectives and indicator selection, with influence on financial allocations being more limited. The analysis also reveals that evaluation market participants from the EU12 are notably more sceptical of use made of “their” evaluation findings. These patterns raise interesting questions as to the role of factors such as the quality and relevance of evaluation recommendations, quality of programming documents drafted and the willingness of planners to take recommendations on board.

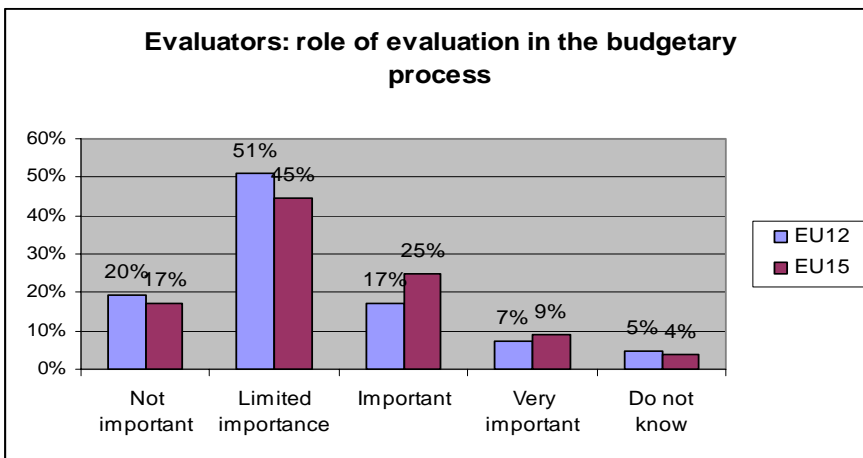
5.1.5. Institutionalisation of Evaluation in the Wider Governance System

The fourth dimension of evaluation capacity benchmarking framework concerns the wider governance environment which affects the development of evaluation practice. Both administrators and evaluation market participants were asked to provide their assessment about the wider governance system and the institutionalisation of evaluation in their country.

Survey results suggest that the role of evaluation in the *national budgetary process* is in many countries of limited importance although the degree of importance was judged to be higher in the EU15. Evaluators’ assessment of the role evaluation plays in the national budgetary process, though slightly more critical, is broadly in line with that provided by administrators.

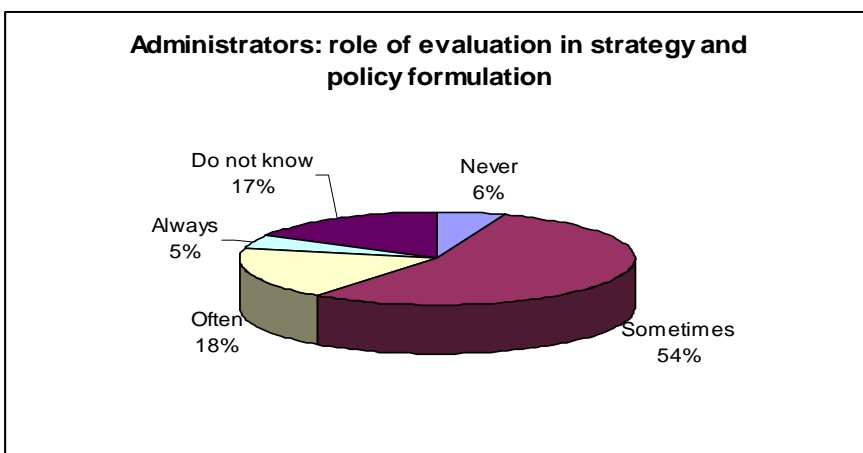


n=108

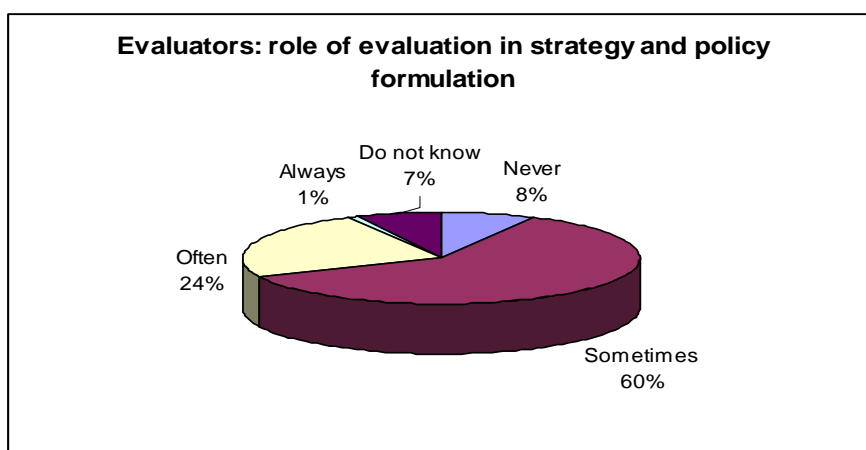


n=117

A similar state of affairs, according to administrators, is apparent as regards the use of evaluations *in strategy and policy formulation* at sectoral/ministry level with a majority of respondents of the view that evaluations were only sometimes used in policy formulation. The assessment provided by evaluation market participants is again similar with more than 60 percent of respondents indicating that evaluations were sometimes used in strategy and policy formulation at sectoral/ministry level.



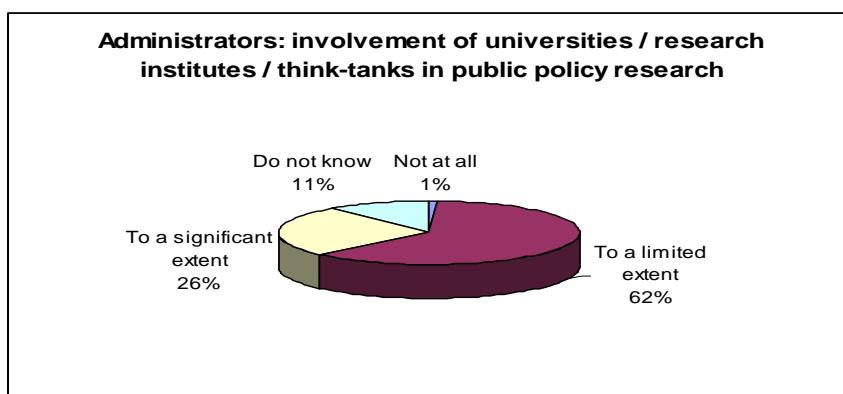
n=108



n=117

Administrators were also asked about the existence of a legal basis setting out *mandatory requirements for the evaluation of national programmes*. However, the results do not allow for any conclusions as in most cases the responses were either contradictory or of a “do not know” form (almost 30 percent). Affirmative answers to this question were received in respect of just three Member States (Germany, Ireland and Slovenia).

A majority of administrators indicated that *research institutions were involved in public policy research* only to a limited extent.



n=108

Asked about the membership in professional evaluation bodies, one third of surveyed market participants indicated they were members of a national professional body of evaluators (37 percent of those from the EU15 and 12 percent of those from the EU12). While responses as regards participation in national and foreign evaluation events are disperse, most frequently evaluators have participated in some 2 national events and in up to 3 events abroad. The total range of national events is 0-20 and those abroad – 0-16. The table 15 below provides a summary of responses to this question. While no universal conclusions can be drawn of the basis of this data of a very “individual” nature, it can be noted that where the specified number of evaluation events at home is high this is an indication that an evaluation society is active.

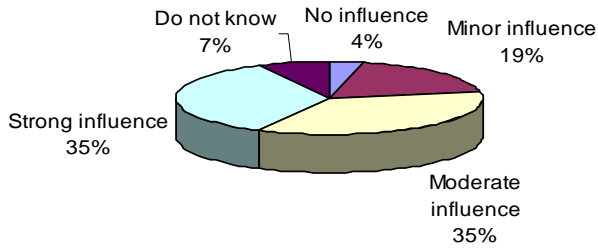
Table 15. National evaluation society

EU Member State	Participation in the evaluation events at home	Participation in the evaluation events abroad	Membership in national evaluation society
Austria	8	2	n.a.
Belgium	2-8	0-3	Belgium evaluation society
Bulgaria	9	8	n.a.
Cyprus	n.a.	n.a.	n.a.
Czech Republic	0-20	0-4	No national professional

EU Member State	Participation in the evaluation events at home	Participation in the evaluation events abroad	Membership in national evaluation society
			body of evaluators
Denmark	10	5	n.a.
Estonia	0-2	0-4	No national professional body of evaluators
Finland	0-20	0-3	n.a.
France	0-10	1-16	French evaluation society
Germany	2-15	1-6	<ul style="list-style-type: none"> ▪ DeGEval - Gesellschaft für Evaluation ▪ Associations of ESF Evaluators DeGEval
Greece	2-18	0-2	No national professional body of evaluators
Hungary	3-10	0-2	In preparation
Ireland	0-20	0-4	Irish evaluation network
Italy	2-5	1-3	Italian evaluation society (AIV Associazione Italiana Valutazione)
Latvia	1	0-1	n.a.
Lithuania	2-5	0-1	n.a.
Luxembourg	1-4	1-5	n.a.
Malta	2	3	n.a.
Netherlands	0-12	1-13	VIDE (Occupational union in the field of evaluation in the Netherlands)
Poland	2-20	0-2	Polish Evaluation Society
Portugal	2-15	1-4	No national professional body of evaluators
Romania	1-5	0-1	Romanian evaluation society
Slovakia	1	0	n.a.
Slovenia	0	2	No national professional body of evaluators
Spain	0-10	1-5	n.a.
Sweden	2-5	2-4	Swedish evaluation society
United Kingdom	0-12	0-10	UK evaluation society

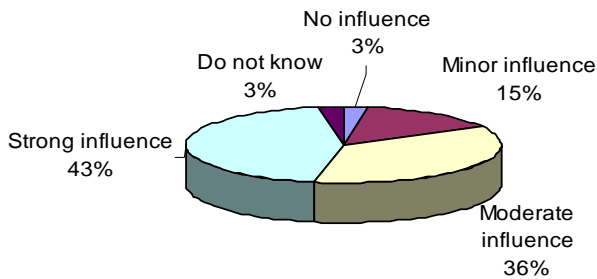
Asked about the *influence of the EU Structural and Cohesion Fund evaluation requirements on national evaluation practice and capacity*, a majority of administrators indicated that they observed such an influence with views split evenly as regards whether the degree of influence was strong or moderate. Similar opinions on the influence of the EU Structural and Cohesion Fund evaluation requirements were expressed by market participants.

Administrators: influence of EU requirements on evaluation practice and capacity



n=108

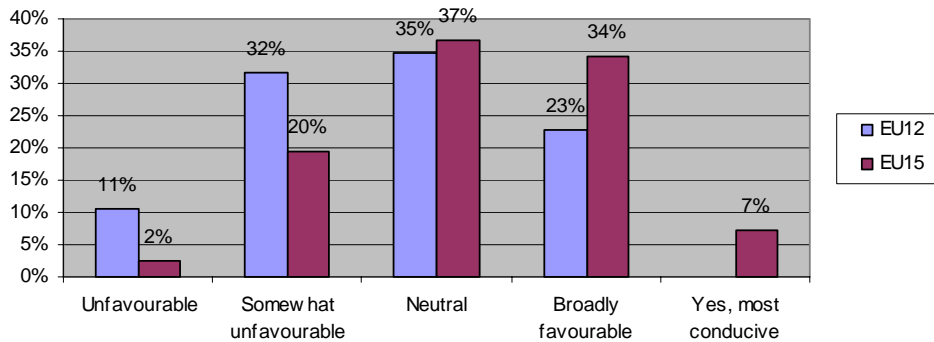
Evaluators: influence of EU requirements on evaluation practice and capacity



n=117

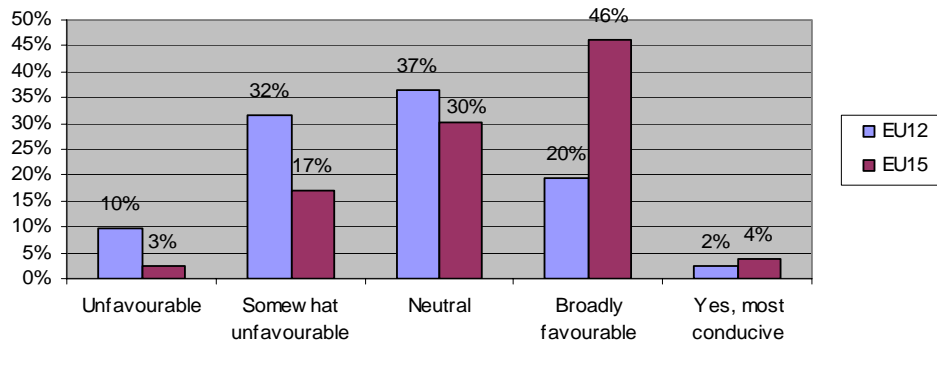
Survey results concerning the *effects of overall administrative and political culture* in their country on the use of evaluation in public policy-making suggest that there is no clear direct effect. Administrators from the EU15 tend to think that the overall administrative and political culture in their country is neutral or broadly favourable to evaluation, while their counterparts from the EU12 see overall administrative and political culture as neutral or somewhat unfavourable to evaluation in public policy making. The opinion of evaluation market participants is broadly in line with that of administrators.

Administrators: effect of administrative and political culture on use of evaluation



n=108

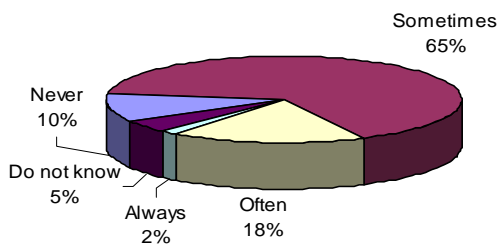
Evaluators: effects of administrative and political culture on use of evaluation



n=117

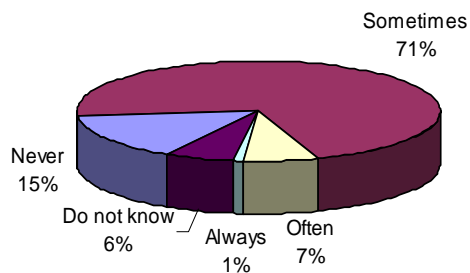
Both groups of respondents were asked about utilisation of evaluation reports and findings *in wider public debates*. A majority of both administrators and evaluators thought that evaluation reports and findings were *sometimes* used in wider public debates. However, only 18 percent thought that reports were *often* used in this context.

Administrators: utilization of evaluation reports and findings in wider public debates



n=108

Evaluators: utilization of evaluation reports and findings in wider public debates



n=117

Summary of findings concerning wider governance environment

Summing up analysis of institutionalisation of evaluation in the wider governance system, the following observations can be made.

- Both administrators and evaluation market participants agree that the role of evaluation in the national budgetary and policy-making process is of limited importance.
- Administrators from the EU15 tend to think that overall administrative and political culture in their country is neutral or broadly favourable to evaluation, while their counterparts from the EU12 are more critical and see administrative and political culture as neutral or somewhat unfavourable to evaluation in public policy making. The opinion of evaluation market participants is similar.
- Administrators and market participants agree that EU Structural and Cohesion Fund evaluation requirements had influence on evaluation practice and capacity in their country, with views on the extent of this influence ranging from “moderate” to “strong” in most cases.

5.2. Methods for Application of the Benchmarking Framework: Lessons Learnt

The previous section outlined the fieldwork findings with regard to evaluation capacity development in the EU Member States. A benchmarking of individual EU Member States was not possible during a piloting exercise, even though interesting data was collected and insights offered on a general level. A comprehensive analysis and comparison was undermined by data deficiencies. In many instances this is not because of irrelevant benchmarking indicators but rather a function of the method of data collection. The testing exercise was useful in that it revealed the kind of data that could be collected via different methods. In the section below, issues around the feasibility of data collection are considered in advance of a consideration of revisions to the benchmarking framework.

5.2.1. Application of the Benchmarking Framework Through Survey

The survey was a useful exercise in several respects. It enabled the identification of the current state of Structural and Cohesion Fund evaluation capacity development, provided insights regarding different evaluation capacity dimensions and allowed for a comparison of views of Structural and Cohesion Fund administrators and evaluation market participants. Survey evidence suggests that on a number of indicators differences exist between the EU15 and EU12. (Exploration of other cleavages, for example, between Northern and Southern countries, as often referred to in literature, was undermined by response rate variations). However, the exercise also pointed-up significant limitations in the application of the proposed benchmarking framework through a survey approach.

Collecting factual information

The collection of quantitative indicators via a survey was of limited use. It had been expected that against this data comparisons based on facts (rather than views) would be possible, e.g. how EU Member States differ in their approaches to organising and managing the evaluation function, in measures taken to develop evaluation capacity and the like. In this regard, the key finding is that collecting so called “hard” data through survey is difficult. Because respondents have different information and levels of awareness about certain issues, the availability of accurate data is problematical. Problems arise immediately when contradictory responses are provided by respondents from the same country or institution. This was the case in respect of virtually all the factual questions in our survey. Relatively straightforward “yes and no” questions proved to be of limited use as both responses generally featured for the same country or institution. Where figures

were asked for, the numbers provided would require cross-checking against other sources – a suspicion confirmed when carrying case studies. Open-ended questions proved to be more useful, but they would require considerable effort in standardising the data collected. In addition, there are limits to how many open questions can be included in a survey given the importance of avoiding an excessively demanding burden on the respondents. Against this background, a conclusion can be drawn that a survey is not an appropriate method to collect data against quantitative indicators.

Collecting opinion-based information

As regards opinion-based questions, the variance in questionnaire return rate per country hindered rigorous cross-comparison of EU Member States. For this type of question representative answers are crucial to be able to infer valid generalisations and conclusions. Thus whilst interesting insights could be made at the target group level or concerning the EU15 and EU12 groupings, little country-based analysis and comparison could be made.

The question arises as to whether it is worth undertaking a survey on the opinion-based questions (or qualitative benchmarking indicators) in future under the condition that representativeness of answers is ensured. For one thing, it could only be a complementary method as opinion-based information alone cannot provide a basis for benchmarking. Moreover, only a very limited number of opinion-based questions could be put to Structural and Cohesion Fund administrations on issues such as, for instance, the adequacy of financial and human resources for evaluation, assessment of capacity and skills, and adequacy and quality of external supply of evaluation services. (As discussed later, some of the opinion-based questions under the fourth dimension of the benchmarking framework would be better addressed to a different target group outside the Structural and Cohesion Fund administrations). Thus it is questionable whether it is worthwhile launching a cross-national survey covering only a few questions, feedback on which would serve as supplementary information. This could only be recommended on a national level if a self-assessment is undertaken. As regards the evaluation market participants, the conduct of an opinion survey would face the significant challenge of defining a target group of evaluators across the EU.

Defining survey target group

An additional lesson drawn from survey exercise concerns survey target group. Surveying programme level institutions is challenging especially as regards factual questions. The problem is that arrangements and procedures below national level may be programme-specific, and answers from some programme level institutions do not allow general conclusions to be drawn.

Against this background, it is clear that benchmarking of Structural and Cohesion Fund evaluation capacity in the EU Member States by way of survey is a challenging exercise and of doubtful cost-effectiveness in terms of what can be collected. The following section considers the appropriateness of an alternative case study method for application of the benchmarking framework.

5.2.2. Application of the Benchmarking Framework Through Case Study

Within the framework of this project, the survey on evaluation capacity development was complemented by four country case studies. This was intended to further test the proposed evaluation capacity benchmarking framework and to see to what extent data against the different benchmarking indicators can be collected through this method.

Case studies yielded a positive assessment on the relevance and applicability of the benchmarking framework. The key finding from the case study experience is that the framework is applicable to different countries and that comparable data can be collected against benchmarking indicators. On

the methodological side, the framework provided a good basis for structuring the fieldwork and for the subsequent collection and analysis of information on evaluation capacity. The application of the benchmarking framework through the case studies allowed for the identification of good practices (something which is much more difficult to do through a survey), and strengths and weaknesses of Structural and Cohesion Fund evaluation capacity in each of the countries. In contrast to the survey, the case studies facilitated contextualised information to be presented and qualitative aspects of evaluation capacity explored. Experience gained in the last programming period could be overviewed and arrangements for the new period presented.

Having said that, some difficulties encountered in the application of the benchmarking framework can be noted. Data on some of the indicators (especially financial resources for evaluation) was not readily available. However, the lack of available information does not suggest that respective indicators should be dropped, but rather this can be seen as an indication of the evaluation management situation in the country.

While case studies provided a good means of addressing qualitative dimensions of evaluation capacity, the assessment of issues such as the quality of terms of reference for evaluation assignments was not straightforward. Careful assessment of this issue would require analysis of a number terms of reference and comparisons either across different institutions or over time. However this would require resources. The same applies to the issue of the quality of evaluation services and expertise.

Another difficulty encountered in carrying out one of the case studies, but not specifically related to the application of the benchmarking framework, was the limited availability of documents and literature in English. This meant that some information collected during interviews could not be effectively cross-checked with other sources and thus could not be used. Therefore, future case studies would benefit from the involvement of experts with proficiency in the language of the Member State concerned.

In conclusion, the application of the evaluation capacity benchmarking framework via case studies was a positive experience which demonstrated that most of the proposed evaluation capacity benchmarking indicators are manageable and that data against them can be collected and analysed. The result of this exercise is a largely comparable set of Member State level reports (attached in Annex 2). The exercise also demonstrates that a case study is a more appropriate method to collect data against evaluation capacity benchmarking criteria and indicators than a survey approach.

5.2.3. Lessons Learnt

All told, careful consideration has to be given to information needs, the appropriateness of different methods for the collection of evaluation capacity benchmarking indicator data and resource availability. In the light of the experience gained, the following conclusions and recommendations for the future can be presented.

- The results of the testing exercise demonstrate that the way to apply the evaluation capacity benchmarking framework is largely through a case study approach, backed up with documentary analysis, rather than via a survey-based method. This would allow for more precision and also a richer, qualitative assessment which would facilitate an examination of how certain formal arrangements function in practice and the identification of good practice examples.
- Country case studies would benefit from the involvement of local experts; this would *inter alia* facilitate review of relevant documentation and literature which may be available only in national language.
- Those responsible for the Structural and Cohesion Fund monitoring function should be interviewed for the purposes of case studies. This would allow for a more specific assessment of monitoring systems. Similarly, case studies would benefit from interviews with people from

outside the Structural and Cohesion Fund administrations in order to explore issues which fall under the fourth dimension of the evaluation capacity benchmarking framework. Information for the assessment of various indicators under the fourth dimension of the evaluation capacity benchmarking framework (namely institutionalisation of evaluation within wider governance), can most effectively be obtained from observers (or documentation) external to the Structural and Cohesion Fund system. Thus case studies can be seen as the most effective method to address this dimension of evaluation capacity development.

- Another approach to collecting data against quantitative indicators (mostly those under the demand side dimension and dissemination and utilisation dimension) could be official correspondence between the European Commission and/or national/regional authorities.

5.3. Implications for the Benchmarking Framework

5.3.1. Strengths and Weaknesses of the Benchmarking Framework

In the light of data collected and experience gained during the piloting exercise, this section considers whether any revisions in the benchmarking framework itself (and not only application methods) are needed to ensure that it is a useful tool to analyse and compare Structural and Cohesion Fund evaluation capacity development in the Member States.

Testing the proposed evaluation capacity benchmarking framework revealed its strengths, weaknesses and necessary revisions. As regards *strengths*, the *benchmarking framework* turned out to be a useful tool in several respects:

- It offers a *comprehensive* approach to evaluation capacity development; it takes into account a wide range of relevant factors and points at multiple facets that need to be considered and addressed for the development of the complex phenomenon that is evaluation capacity.
- It is *sensitive to national contexts and adaptable* to the circumstances of different Member States. It can be used both at a national/regional level and for cross-national comparison. Pilot application of the benchmarking framework allowed the identification of broad tendencies and provided interesting insights into evaluation capacity development among Member States. The application of the framework via case studies of selected countries gave a detailed, contextualised, but yet largely comparable view with good practices and learning points established. If applied regularly and via appropriate data collection methods, it could generate comparative information across Member States at different time intervals – something which has been lacking to date.
- From a methodological perspective, the framework facilitates *systematic* and *logical* analysis of evaluation capacity.

The testing exercise revealed certain *weaknesses* of the proposed evaluation capacity benchmarking framework.

- The comprehensiveness of the framework represents a limiting factor in some respects. It means that the framework is *demanding in information and the resources* to collect it. As will be discussed below, the piloting phase has identified some indicators that can be dropped, thereby partially addressing this concern.
- The piloting experience leads to the conclusion that the benchmarking framework should be seen as *(self)assessment or diagnostic tool rather than a benchmarking tool* in a strict sense. Developing a respective benchmark, standard or target for each indicator is hardly feasible and appropriate, and Member States can hardly be “ranked” on the basis of performance against all indicators even when sufficient data is collected. This indicates a general limitation of applying a benchmarking “technique” in Structural Funds evaluation, which operates within specific

constitutional, institutional and administrative structures. Having said that, it must be emphasised that the proposed evaluation capacity framework is a useful means to monitor and compare evaluation capacity development across the EU, identify areas in need of improvement, good practices and learning points for further evaluation capacity building.

- In the light of testing experience, the *revision of certain indicators* can be recommended and is proposed in the section immediately below. While none of the indicators can be seen as irrelevant, some of them can be dropped or further specified for clarity purposes taking into account data collection feasibility.

5.3.2. Revision of Evaluation Capacity Benchmarking Framework

The demand side dimension

Architecture (Integration, location and organisation of evaluation function)

Most of indicators under this criterion proved both necessary and useful. They constitute a starting point in any analysis of evaluation capacity in particular as regards establishing the location and organisation of the evaluation function within the Structural and Cohesion Fund administrations. An analysis of these indicators reveals differences in the approaches to and degree of institutionalisation of evaluation function in the EU Member States.

- In the light of testing experience, one indicator under this criterion can be omitted – namely, the indicator on the *existence of document delineating evaluation roles and responsibilities*. As a general rule, programming documents outline key evaluation roles and responsibilities; so this indicator tells little new information and is hardly a basis for comparison. In addition, it does not tell how responsibilities are shared among institutions. This particular issue can be adequately explored under the rest of indicators under the “architecture” criterion.

Evaluation resources

All indicators under this criterion are relatively straightforward and indispensable for the assessment and comparison of evaluation capacity. For instance, even with no definite figures provided, application of these indicators identified areas where closer attention is needed (i.e. insufficient human resources for evaluation and lack of readily available information on financial resources). The criterion and its respective indicators should be kept. No revisions emerge from the testing experience.

Efficiency of evaluation management

Taking into consideration fieldwork experience, the following modifications in individual indicators under the evaluation management criterion are suggested.

- The indicator on the *existence of evaluation plans* is further specified to reflect the fact that evaluation plans may be multi-annual and/or annual.
- *The sub-indicator referring to core terms of reference or mandatory evaluation questions* could be dropped. The relevant indicator on the terms of reference requires an assessment of their quality. Moreover, there is a danger that core terms of reference or standard evaluation questions might be applied in a mechanistic manner (although we accept that such an approach may be necessary in situations where evaluation capacity and experience is limited or where, for example, regional or programme level reports constitute inputs to higher, national-level evaluation exercises).
- *Procurement* is an important criterion: the way in which evaluations are commissioned may influence the development of the supply side or utilisation of evaluation results. At the same

time it is a complex issue. In addition to the duration of the procurement process, other aspects – such as contract award criteria (e.g. price only or price/quality ratio) – also play an important role and are therefore worth considering. Observations as regards dominant procurement practice and/or differences between institutions may be expected to yield interesting insights. This is reflected in the revised framework.

- The indicator on the *existence of evaluation steering committees* could be further refined to include such issues as their *composition and key functions*. As fieldwork findings suggest, steering groups are a common practice, and differences exist not so much with regard to existence/non-existence of steering groups at various levels, but in terms of what functions they perform and their composition.

Quality of monitoring systems

The availability of reliable and timely monitoring data is a precondition for good evaluation. Thus quality of monitoring data is an important issue to consider when assessing evaluation capacity. Framework testing experience suggests that virtually no revisions, except one minor adjustment, are necessary.

- Indicator “*quality and reliability of indicator data*” could be rephrased and put as follows: “*quality of indicator data at programme / national level*”

The supply side dimension

Quality of socio-economic data resources

The indicator on the availability of socio-economic indicator data proved useful in that its application revealed that data at regional level is perceived as less accessible than data at national level.

- Conversely, in the light of fieldwork findings, the sub-criterion of the *timeliness of socio-economic indicator data* turns out to be of limited use and could be omitted. On the one hand, EU level (Eurostat) requirements exist governing coverage and timeliness of key statistical data with which EU Member States are expected to comply. In addition, data for different type of socio-economic indicators becomes available at different intervals, thus “*latest year for which data is currently available*” might vary across different categories (e.g., inflation and GDP). Such a detailed consideration would be of limited use.

Availability and quality of evaluation expertise

Indicators under this particular criterion addressed key aspects of evaluation supply. Analysing availability of evaluation supply on the domestic market is relatively straightforward, thus only one minor amendment is suggested in this regard. Assessment of thematic and methodological expertise is highly relevant, but rather demanding. One of the most appropriate methods to address this issue in detail would be through meta-evaluations.

- For simplicity purposes, the indicator on the *number of firms active in the market* and the one concerning the *involvement of universities in evaluation practice* could be merged and expressed as follows: “*number and type of organisations (firms and academic/research institutions) active in the market*”.

Dissemination and utilisation dimension

The importance of dissemination of evaluation outputs cannot be overrated. Experience of the application of the indicator *on public access to evaluation outputs* within the framework of this project allows us to conclude that it faces little difficulty and thus no revisions are suggested. The

utilisation criterion is even more important - it is the basis for making judgements on whether evaluations „matter“ or are more of a formal routine. The piloting exercise carried out within the framework of this project provided interesting insights with regard to impacts of the most recent evaluations. Against this background, only one minimal revision under this dimension is suggested.

- Indicator on the *degree of influence of evaluation on ongoing programme management and implementation (2000-2006)* should be adjusted by eliminating the reference to the 2000-2006 period and in this way making the indicator useful going forward.

An enabling governance environment

Indicators under this dimension take into account the national context for Structural and Cohesion Fund evaluations, and also serve as the real "test" of whether evaluation is taken seriously and whether it is likely to be sustainable. In this regard indicators on the use of evaluation in the national budgetary and policy-making processes are of particular relevance. However, the overall number of indicators under this dimension could be reduced for the sake of simplicity and feasibility of data collection.

- The indicator on the *existence of programme of public management reform* could be dropped as this issue, if relevant, can be covered under indicators on the use of evaluation in national budgetary process and policy-making.
- In addition, based on the framework testing experience, the indicator on the *existence of public policy research bodies or think-tanks* provides little information and is therefore of limited use and could be omitted.
- At the same time one new indicator is suggested under the criterion relating to evaluation society: i.e. *key activities of evaluation society*. It is of interest to explore by what specific means and activities evaluation societies contribute to the development of evaluation capacity, and not just establish figures concerning membership or evaluation events. This new indicator would draw attention to this issue.

All told, it is clear from the discussion above that the essence and structure of the proposed evaluation capacity development framework is maintained and only limited revisions are made taking into account the experience gained during the piloting exercise. The number and content of evaluation capacity dimensions and criteria remains unaltered. The number of sub-criteria is reduced to 25 and the number of indicators to 35. The revised version of framework is presented in the table below. The table incorporates not only suggested revisions, but also recommended data collection methods for each indicator, which should facilitate any possible future application of the framework.

**Revised Framework to Analyse Evaluation Capacity Development:
Dimensions, Criteria, Sub-criteria and Indicators**

Dimension	Criterion	Sub-criteria	Indicators	Recommended data collection method
DEMAND SIDE	Architecture (Integration, location and organisation of Structural and Cohesion Fund evaluation function)	Evaluation Responsibility	1. Location of responsibility for evaluation function within the managing authority/authorities	Official correspondence <i>or</i> case study
			2. Existence of dedicated evaluation function (Evaluation Units / national and programme level)	
		Coordination	3. Existence of formal mechanisms for evaluation coordination across programmes	
		Links with other functions	4. Linkages between evaluation function and other related functions, such as monitoring and programming	Case study
	Evaluation Resources	Financial	5. Evaluation budgets (euros)	Official correspondence <i>or</i> case study
		Staffing	6. Number of evaluation staff (national and programme level)	Case study (interview/ desk research)
			7. Evaluation experience and skills of evaluation staff <ul style="list-style-type: none"> ▪ Evaluation commissioning experience ▪ Evaluation practice experience 	
		Evaluation flow and coverage	8. Number of evaluations carried out / time period	Official correspondence <i>or</i> case study
	Efficiency of evaluation management	Planning	9. Existence of multi-annual and annual evaluation plans at national and programme level	Official correspondence <i>or</i> case study
		Terms of Reference	10. Quality of evaluation Terms of Reference	Case study
		Procurement	11. Average interval between agreement on TOR and appointment of evaluator	
			12. Contract award criteria	
		Evaluation Steering Committees	13. Existence, composition and functions of evaluation steering committees (national and programme level)	Case study <i>or</i> official correspondence
		Quality Assessment	14. Existence of approved set of quality standards	

Dimension	Criterion	Sub-criteria	Indicators	Recommended data collection method
	Quality of Monitoring Systems	Reliability	15. Quality of indicator data at national /programme level	Case study
		Timeliness	16. Average interval between year end and availability of indicator data for year in question	
SUPPLY SIDE	Availability of socio-economic data	Availability	17. Availability of key socio-economic indicator data (GDP, employment, unemployment, R&D investment) at regional (NUTS II) level	Case study
	Availability and quality of evaluation expertise	Availability of evaluation supply on domestic market	18. Number and type of organisations active in the market	Case study
			19. Existence of internal evaluation unit capacity	Official correspondence <i>or</i> case study
			20. Availability of evaluation training/education options	Case study <i>or</i> official correspondence
		Thematic and methodological expertise	21. Breadth and quality of expertise (areas where expertise is particularly strong or deficient: e.g., macroeconomic modelling, horizontal principles, infrastructure, human resources etc.)	Case study / meta-evaluation
			22. Quality of evaluation reports	Case study / meta-evaluation
DISSEMINATION AND UTILISATION	Dissemination	Access to evaluation outputs	23. Public access to evaluation reports	Case study <i>or</i> official correspondence
	Utilisation	Procedures for addressing evaluation and follow-up procedures	24. Existence of formal monitoring committee or managing authority responsibilities	Official correspondence <i>or</i> case study
			25. Existence of follow-up procedures where evaluation recommendations are agreed	Official correspondence <i>or</i> case study
		Impacts on programme design and implementation	26. Degree of influence of ex ante evaluation of 2007-2013 period on: <ul style="list-style-type: none"> ▪ Strategy formulation and programme design ▪ Allocation of resources ▪ Quantification of objectives and indicator selection 	Case study
		27. Degree of influence of evaluation(s) on ongoing programme management and implementation	Case study	

Dimension	Criterion	Sub-criteria	Indicators	Recommended data collection method
INSTITUTIONALISATION OF EVALUATION IN THE WIDER GOVERNANCE SYSTEM	An enabling governance environment? (Degree of institutionalisation of evaluation in the wider governance system)	Role of Evaluation in State Budget/degree of integration of evaluation into budget process	28. Degree to which evaluations represent routine part of budgetary process	Case study
		Role of evaluation in sectoral strategy/policymaking	29. Extent to which evaluations are routinely used in strategy and policy formulation at sectoral/ministry level	Case study
		Existence of mandatory evaluation requirements	30. Existence of statutory evaluation requirements	Case study (desk research)
		Existence of an evaluation society or network	31. Membership size	Case study
			32. Frequency of meetings/conferences	
			33. Key activities	
		Influence of EU evaluation requirements	34. Degree of influence of EU Structural Funds evaluation requirements on Member State evaluation practice	Case study
Role of civil society and other players	35. Extent to which evaluation findings are drawn on by civil society and social partners	Case study		

6. CONCLUSIONS AND RECOMMENDATIONS

Evaluation is a core component of the Structural and Cohesion Fund programming and management processes. Evaluation plays a role in the design of new programmes, in supporting the implementation process and in assessing the effectiveness of interventions. For the 2007-2013 programming period important modifications to the Structural and Cohesion Fund evaluation approach have been introduced aimed at better integrating evaluation into programme management and implementation. Going forward, ongoing evaluation will be designed and undertaken in accordance with internal demands at the Member State level. This requires not only strong links between monitoring and evaluation, but adequate evaluation capacity generally to ensure the quality of on-going evaluation and its contribution to decision-making.

The development of evaluation capacity is a multi-dimensional process that requires time, patience and persistence. The literature on the subject suggests that efforts to develop evaluation capacity benefit from a diagnosis and regular monitoring of evaluation systems and practice to detect what is working and areas requiring attention. In addition, experience of other countries and good practice examples can serve as sources of inspiration and learning.

So far systematic and comparable information on Structural and Cohesion Fund evaluation capacity and its development in the EU Member States has been lacking. This research attempted to address this issue by developing a tool to assess and compare evaluation capacity development within Structural and Cohesion Fund administrations. The final chapter herein presents overall conclusions which arise from this exercise. While very specific recommendations on actions to be taken at various levels can hardly be provided given the testing nature of the exercise, a number of recommendations emerge from the analysis. They concern both: (i) methodological issues around the future application and use of the framework; and (ii) substantive issues concerned with evaluation capacity development within Structural Fund administrations. Recommendations are addressed to different stakeholders.

6.1. On the Framework to Analyse Evaluation Capacity Development

1. ***Limits of benchmarking.*** The initial intention of developing a framework to analyse evaluation capacity development was to produce a benchmarking tool. However, the experience gained demonstrates that setting benchmarks and ranking performance is hardly feasible and appropriate in the field of Structural Fund evaluation, where complex institutional structures, wider administrative environment and historical and cultural contexts need to be taken into account. There is no one best practice approach to Structural and Cohesion Fund evaluation capacity development.
2. ***A flexible analytical framework.*** The research resulted in a tested and ready-to-use analytical framework that systematically addresses multiple aspects of evaluation capacity development and allows for a review of progress in this field. The framework has been informed by a literature review, a consideration of the key features of the Structural and Cohesion Funds evaluation model and the results of the piloting exercise. Going forward, the framework to analyse evaluation capacity development can be used as:
 - (a) a *means for monitoring* and comparing *evaluation capacity development across the EU*, identifying trends, good practices, learning points and areas in need of improvement;
 - (b) a *self-assessment or diagnostic tool* by any Member State willing to undertake an assessment of Structural and Cohesion Fund evaluation capacity at a country/region

level. The various criteria and indicators in the framework highlight different aspects of evaluation capacity and allow areas in need of further improvement to be identified.

3. **Methods of application.** The most appropriate method to apply the framework is largely through a case study approach with some local input and, on some dimensions, official correspondence.

Recommendations

- For the European Commission**
- Present the framework to analyse evaluation capacity development to Member States and encourage them to use it for (regular) analysis of their own situation (for instance, for evaluation plan purposes).
 - Use the framework to analyse evaluation capacity development in the upcoming ex post evaluation of Structural Fund management systems (namely, concerning evaluation systems).
 - Collect data at regular intervals against indicators for which official communication means can be used (mostly those under the demand side and dissemination and utilisation dimensions).
 - Consider undertaking (commissioning) more in-depth analysis of evaluation capacity development by applying the framework via case studies. In such instances, it is recommended to carry out analyses of different Member States at the same time in order to produce data which is comparable in time and across the EU.
 - Disseminate identified good practice lessons in evaluation capacity development.
- For Member States**
- Member States are encouraged to apply the framework for their self-assessment. Evaluation capacity analyses are recommended and could be included as projects in their evaluation plans. Such a self-assessment of evaluation capacity is desirable because it can guide the identification of strengths, weaknesses and opportunities for further improvement.

6.2. On Evaluation Capacity Development in the EU Member States

The demand side

1. **Institutionalising evaluation.** The dominant practice in the Member States is to set up formal structures for the evaluation of Structural and Cohesion Funds. Specialised evaluation units operate within managing authorities in a number of Member States. Steering groups and committees to coordinate and manage evaluation issues have been set up in almost all countries. The development of evaluation plans is becoming a prevalent practice in the 2007-2013 programming period; this is a positive feature and should assist Member States in managing evaluation demand in the 2007-2013 period.
2. **Human resources not yet adequate.** The dominant opinion of Structural and Cohesion Fund administrations is that financial allocations for evaluation are generally adequate but that human resources are not, particularly in the EU12. Without formal requirements for evaluation, a sufficient degree of ownership in Member States will be needed to ensure that evaluation is taken seriously. Where institutions are understaffed or where staff are occupied with other functions besides evaluation, it is much more difficult to analyse and plan evaluation needs,

which represents a risk factor in the context of the new system of on-going evaluation. This also limits the possibility of time and resources being devoted to the development of specific expertise and skills in evaluation field. In addition, survey evidence suggests that supply of training options is limited in a number of Member States, especially the EU12.

3. ***Different levels of awareness.*** National authorities (especially at programme level) tend to have different information and different level of awareness about various aspects of Structural and Cohesion Fund evaluation, a finding which further confirms the value of undertaking analysis of evaluation capacity and disseminating its results.
4. ***Monitoring systems need further improvement.*** The timeliness and quality of monitoring indicator data is still a weak area in a number of countries, especially in the EU12. This can be seen as a risk as well-functioning monitoring systems and the availability of high quality timely data are pre-requisites for good evaluation, especially in the context of on-going evaluation systems.

Recommendations for strengthening the demand side

For the European Commission

- The recommendation which emerges from the interviews with national authorities concerns the value of organising *specific themed workshops at the EU level* which would discuss very practical evaluation issues through a combination of presentations from prominent experts in the field and exchange of Member States' experience. Authorities feel that such issue-specific events are lacking and other training of more general nature cannot fill this gap.

For Member States

- Targeted efforts at strengthening human resources are necessary. *Training* and other evaluation capacity building measures remain crucial if the efficiency of evaluation management is to be increased and if capacity to carry out internal and external evaluations is to be further developed. For this reason, it is recommended to assess, plan and include evaluation capacity building activities in the forthcoming evaluation plans. In addition, information about existing and available training opportunities needs to be better circulated within administrations. This information could be disseminated via institutional websites (sections dedicated to evaluation), intranets, evaluation steering groups, etc.
- Apart from training, *periodic events on exchange of experience* between different national institutions would also contribute to further evaluation capacity development within administration and higher awareness of various aspects of the entire system for evaluation. Such events could focus on such issues as experience gained in commissioning evaluations (procedures, steering experience, quality of reports, impact, lessons learnt, etc.) or carrying internal evaluations (where applicable).
- *Evaluations to assess the quality of monitoring systems and data* are recommended. There are already initiatives undertaken in some countries to evaluate the functioning of monitoring systems with the aim of improving them. This is a good practice example that could be considered by other Member States.

The supply side

1. ***External supply of evaluation services.*** In some Member States evaluation units carry out internal evaluations, but in most cases the supply of evaluation services remains external to the

administrations. In many countries, local private companies tend to dominate the provision of services, with limited involvement of research institutions or foreign expertise. Internal evaluations could be useful for some types of evaluations, in particular in the context of on-going evaluations, but limited capacity still exists in this regards, especially in the EU12.

2. **Differences in evaluation market development.** Research evidence supports the widely-held opinion that evaluation markets in many of the EU12 countries are still under-developed.
3. **Still room for improvement in quality of evaluation services.** While the general opinion - as expressed by both evaluators and administrations - on the quality of external evaluation services is quite positive, there is certainly room for improvement for instance, in areas such as quantitative and experimental techniques, horizontal issues, etc. Training and education opportunities for evaluators are still limited in the EU12.

Recommendations for strengthening the supply side

For Member States It is more difficult to come up with recommendations to strengthen the supply side since it develops outside the administrative system. However, the following means to reinforce evaluation capacity development on the supply side could be considered.

- The commissioning of *more evaluations* would provide more possibilities for market participants to take part in evaluation activities and thereby contribute to strengthening capacity. There needs to be sufficient demand for evaluation if market participants are to specialise in the provision of such services and invest in strengthening of relevant capacities. This is particularly important in the EU12 which have a short track record in evaluation. More *smaller scale thematic* evaluations have the potential to raise interest in providing evaluation services and to attract new players into the market.
- National authorities can create incentives for the development of external evaluation expertise by using *tendering policy and requirements* as an instrument (e.g. attempting to widen evaluation community, attract academia or consortia with foreign expertise).
- Member States are encouraged to *publicise evaluation plans* to inform the wider evaluation community as to the amount of evaluation work foreseen.
- Member States are encouraged to consider undertaking so-called *meta-evaluations* of evaluation reports to assess their quality, methods used, results received, utilisation of recommendations and the like. Such an exercise would constitute a learning experience both for administrations and evaluators (if publicised). It would potentially contribute to better quality control and quality of evaluation reports.
- Advertise *evaluation events* organised by public authorities and *invite evaluators* to participate in them.

For evaluation market participants ▪ Here it could be emphasised that evaluation societies contribute to developing evaluation capacity by providing opportunity for debate and encouraging learning. Thus actions taken to establish evaluation societies in a number of Member States constitute good practice examples.

Dissemination and utilisation

1. **Dissemination quite good.** Evaluation reports are quite widely available to the public, including via internet. Accessibility to reports (or at least awareness of them) is higher in the EU15.

2. **Utilisation of results can be improved.** Member States have different arrangements in place to address evaluation findings with various functions exercised by monitoring committees, evaluation steering groups, evaluation units or managing authorities. However, the assessment of the influence of the most recent ex ante evaluation, which is closely integrated into the programming process, shows that utilisation of evaluation results is still limited. Research evidence suggests that impact of ex ante evaluations was greatest on the quantification of objectives and indicator selection, with influence on financial allocations being the most limited. Research results also suggest that there exists a credibility gap in the EU12 where administrators are more critical of quality of external evaluation services and where evaluators feel that their results are not sufficiently used and disseminated.

Recommendations for strengthening dissemination and utilisation

- For Member States** Although the general availability of evaluation reports is quite good across the EU, there is still room for development in this area, especially in the EU12.
- Such measures as *centralised online access* to Structural Fund evaluation reports can contribute to better awareness of evaluation projects and their higher impact.
 - Some Member States have set *minimum requirements for publication* of the evaluation reports (e.g., requirement that at a minimum, the executive summary of evaluation reports shall be made publicly available). This is a practice which could be considered by other countries.
 - Where these do not exist, it is recommended to consider setting up *formal procedures to address and follow-up of evaluation results* in order to ensure their better integration into decision-making.

Wider governance environment

1. **National evaluation practice in most cases limited.** Administrators and evaluation market participants regard the role of evaluation in the national budgetary and policy-making processes as being of limited importance and note that in most cases evaluation reports are only sometimes used in wider public debates. The assessment of the influence of EU Structural and Cohesion Fund requirements on national evaluation practice and capacity generally ranges from “moderate” to “strong”. Administrators from the EU15 tend to think that the overall administrative and political culture in their country is neutral or broadly favourable to evaluation, while their counterparts from the EU12 are more critical and see administrative and political culture as neutral or somewhat unfavourable to evaluation in public policy making. The opinion of evaluation market participants is similar.

Recommendations in relation to wider governance environment

- For European Commission** ▪ The European Commission should encourage Member States to *use Structural funding for evaluation capacity building* within wider administration via Administrative Capacity OPs (funded by the ESF in a number of Member States) or through establishing specific objectives in the ESF programmes, which focus on the development of institutional capacity.
- For Member States** ▪ *Administrative Capacity OPs* (funded by the ESF in a number of Member States) or ESF programmes could be used as a vehicle *to enhance evaluation capacity* more widely and make the wider administrative culture (outside the scope of Structural Fund system) more conducive to evaluation practice.

BIBLIOGRAPHY

EU Legal Texts

1. Council Regulation No. 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999;
2. Regulation (EC) No 1080/2006 of the European Parliament and of the Council of 5 July 2006 on the European Regional Development Fund and repealing Regulation (EC) No 1783/1999.
3. Regulation (EC) No 1081/2006 of the European Parliament and of the Council of 5 July 2006 on the European Social Fund and repealing Regulation (EC) No 1784/1999.
4. Regulation (EC) No 1082/2006 of the European Parliament and of the Council of 5 July 2006 on a European grouping of territorial cooperation (EGTC)
5. Council Regulation (EC) No 1084/2006 of 11 July 2006 establishing a Cohesion Fund and repealing Regulation (EC) No 1164/94
6. Council regulation (EC) No.1260/1999 of 21 June 1999 laying down general provisions on the Structural Funds

Methodological Guidance

1. European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013. Indicative Guidelines on Evaluation Methods: Ex Ante Evaluation. Working Document No.1. August 2006.. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd1_exante_en.pdf
2. European Commission, Directorate-General for Regional Policy. New Programming Period, 2007 – 2013. Indicative Guidelines on Evaluation Methods: Monitoring and Evaluation Indicators. Working Document No.2. August 2006. // http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd2_indic_en.pdf
3. European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf
4. Commission of the European Communities. The 2000 – 2006 Programming Period: Methodological Working Papers. Working Paper 8. The Mid Term Evaluation of Structural Fund Interventions. Brussels: 5 December 2000.
5. European Commission, Directorate-General for Regional Policy. The 2000 – 2006 Programming Period: Methodological Working Papers. Working Paper 9. The Update of the Mid Term Evaluation of Structural Fund Interventions. http://ec.europa.eu/regional_policy/sources/docoffic/working/doc/midterm_update_en.pdf
6. European Commission (2000), Good Practice Guidelines for the Management of the Evaluation Function, prepared by the Working Group on Guidelines. http://ec.europa.eu/budget/evaluation/goodpractice/goodpractice_en.htm
7. Guidelines from DG Budget http://ec.europa.eu/budget/evaluation/index_en.htm
8. Evaluation in DG EMPL http://ec.europa.eu/employment_social/evaluation/index_en.html

Literature

1. Assessing the Evidence: The Evaluation of Regional Policy in Europe, European Policies Research Centre, 2004
2. Bachtler J., Michie R. „A new era in EU regional policy evaluation: the Appraisal of the Structural Funds”, *Regional Studies*, Vol. 29.8, p.745-751.
3. Bachtler J., Polverari Laura, Sandra Taylor, Brian Ashcroft and Kim Swales, „Methodologies used in the Evaluation of the Effectiveness of European Structural Funds: A Comparative Assessment”, 2000
4. Batterbury, Sarah C., „[Principles and purposes of European Union Cohesion policy evaluation.](#)“ *Regional Studies*, Vol. 40 Issue 2, 2006, p179-188
5. Blažek Jiří, Jan Vozáb, „Forming evaluation capacity and culture in the Czech Republic: experience with the first set of ex ante evaluations of programming documents (with special focus on evaluation of NDP)“, 2003.
http://ec.europa.eu/regional_policy/sources/docconf/budapeval/work/blazek.doc
6. Boyle R., „Evaluation Capacity Development in the Republic of Ireland”, *ECD Working Paper Series* n°14. The World Bank, June 2005.
7. Boyle Richard, Donald Lemaire. Building Effective Evaluation Capacity. Lessons from practice. - Transaction Publishers, 1999 – 202 p.
8. Casavola, Paola and Laura Tagle, „Building capacity for evaluation: Lessons from Italy”, *Fifth European Conference on Evaluation of the Structural Funds Budapest, 26/27 June 2003*
9. Centre for Strategy and Evaluation Services, Benchmarking Evaluation Capacities in New Member States, 2005.
<http://www.cses.co.uk/publications/Benchmarking%20eval%20capacity%20in%20NMS.doc>
10. Compton, D. M. Baizerman and S. Stockdill (Eds) The Art, Craft and Science of Evaluation Capacity Building, New Directions for Evaluation - 93, Spring 2002. San Francisco, CA: Jossey-Bass.
11. Dabelstein, Niels. Evaluation Capacity Development: Lessons Learned // *Evaluation*, Jul 2003; vol. 9: pp. 365-369.
12. Davies, P., 2004. Is Evidence-Based Government Possible? *4th Annual Campbell Collaboration Colloquium*, 19 February 2004 Washington D.C.
13. DG Regional Policy Evaluation Network, „The Mid Term Evaluation in Objective 1 and 2 Regions Growing Evaluation Capacity.“ Final Report, 2004.
http://ec.europa.eu/regional_policy/sources/docgener/evaluation/tech_en.htm
14. Feinstein O., N., 2002. Use of Evaluations and the Evaluation of their Use. *Evaluation*, Vol. 8(4), pp. 433–439.
15. Fifth European Conference on Evaluation of the Structural Funds, „Challenges for Evaluation in an Enlarged Europe”, Budapest, 26/27 June 2003.
http://ec.europa.eu/regional_policy/sources/docconf/budapeval/index_en.htm
16. Forss, Kim, Kruse, Stein-Erik, Taut, Sandy, Tendén, Edle. Chasing a Ghost?: An Essay on Participatory Evaluation and Capacity Development // *Evaluation*, Jan 2006; vol. 12: pp. 128-144
17. Furubo, Rist and Sandahl (eds) International Atlas of Evaluation. - Transaction Publishers, 2002
18. Grant, J. (2005) “Getting the Evidence: using the research in policy making” in *Evidence-based Policy Making Getting the Evidence, Using the Evidence and Evaluating the Outcomes*, Conference Proceedings, Dublin: National Economic and Social Forum.
19. Guerrero R. Pablo O. *Evaluation Capacity Development, Comparative Insights from Columbia, China and Indonesia.* - World Bank Operations Evaluation Department, January 1999
20. Hanberger A., 2001. What is the policy problem? *Evaluation*, Vol. 7 (1) pp. 45-62

21. Hegarty, David, Presentation by to Commission Responsible for NSRF, „Evaluation on EU Programmes in Lithuania”, Vilnius, 8 September 2005.
22. Hegarty, David and Fitz Gerald, John, „Ex Ante Evaluation Process for the 2000-2006 Period in Ireland“. http://ec.europa.eu/regional_policy/sources/docconf/edimbourg/pdf/hegar_en.pdf
23. Hegarty, D, “Framework for the Evaluation of the Structural Funds in Ireland” *Fifth European Conference on Evaluation of the Structural Funds Budapest, 26/27 June 2003* <http://www.csfinfo.com/documents/publications/evaluation/ConferencePaper.doc>
24. Jaszczolt, Krzysztof, Tomasz Potkański and Stanisław Alwasiak, „Internal Project M&E System and Development of Evaluation Capacity – Experience of the World Bank-funded Rural Development Program.” http://ec.europa.eu/regional_policy/sources/docconf/budapeval/work/jaszczolt.doc
25. Khan, M. Adil. Evaluation Capacity Building: An Overview of Current Status, Issues and Options. *Evaluation*, Jul 1998; vol. 4: pp. 310-328
26. Knaap P., “Performance Management and Policy Evaluation in the Netherlands. Towards an Integrated Approach.” *Evaluation*, Vol. 6, No. 3, 2000, p.335-350
27. Lahey, R., *A Framework for Developing a Monitoring and Evaluation System*, Presentation to the International Conference on Monitoring and Evaluation, Kiev, Ukraine: February 2007.
28. Lang, J., 2001. Improving Structural Policy Under Conditions of Hybrid Governance: Multi-Actor Constellations, Integrated Feedback Instruments and the Use of Evaluation Results. *Evaluation*, Vol 7(1), pp. 7–23.
29. Lewis J., 2001. Reflection on the evaluation practice, *Evaluation*, Vol 7 (3) pp.387-394
30. Mackay Keith. Development of Australia’s Evaluation System. - World Bank - ECD Working Paper 1998.
31. Mackay Keith. Evaluation capacity development: a diagnostic guide and action framework. – The World Bank Operations Evaluation Department. – ECD Working Paper series. – No.6. – 1999.
32. Mackay Keith. Institutionalization of Monitoring and Evaluation Systems to Improve Public Sector Management. - The World Bank Operations Evaluation Department - ECD Working Paper Series - No. 15. - January 2006
33. Mackay, K. "The World Bank's ECB Experience" in Donald W. Compton, Michael Baizerman and Stacey Hueftle Stockdill (eds.) *The Art, Craft and Science of Evaluation Capacity Building, New Directions for Evaluation*, 2002
34. Mairate, A. Developing evaluation capacities in the Member States: the case of Structural Funds”, IDEAS Workshop, Prague 19/20 June 2006. <http://www.ideas-int.org/Documents/Mairate%20paper.doc>
35. Malan, Jack, Presentation: „Benchmarking Evaluation Capacities in New Member States”, Building Evaluation Capacity and Culture in the Baltic States, Vilnius, 28/29 November 2005
36. McDonald, Bron, Rogers, Patricia, Kefford, Bruce. Teaching People to Fish? Building the Evaluation Capability of Public Sector Organizations // *Evaluation*, Jan 2003; vol. 9: pp. 9-29
37. Nutley, S. & Davies, H., 2000. Making a reality of evidence-based practice. In: H. DAVIES, S. NUTLEY, P. SMITH ed. *What Works? Evidence-based Policy and Practice in Public Services*. Bristol: The Policy Press, pp.317-351.
38. Picciotto, R. (1998) “Evaluation Capacity Development: Issues and Challenges” in *Evaluation Capacity Development in Africa Selected Proceedings from a Seminar in Abidjan*, Washington, D.C.: Operations Evaluation Department, The World Bank.
39. Picciotto, R. (1999) “Towards an Economics of Evaluation”, *Evaluation* 5(1): 7-22
40. Polverari L., Bachtler J., „Assessing the Evidence: the Evaluation of Regional Policy in Europe”, EoRPA paper 04/5, European Policies Research Centre, September 2004.

41. Ratajczak, Justyna, Beata Cieżka, „Creating the structural funds’ evaluation system in Poland - opportunities for dialogue and cooperation between evaluators and stakeholders” 6th European Evaluation Society Biennale Conference „Governance, Democracy and Evaluation”, 2004
42. Reines Philip and Sandra Taylor, „Mid –term Evaluation of the 2000-06 Structural Funds Programmes”. *IQ-Net Thematic Paper*, 11(2), 2002
43. Sanderson I., 2004. Getting evidence into practice. Perspectives and Rationality. *Evaluation*, Vol. 10(3), pp.366-379.
44. Stern, Elliot. What Shapes European Evaluation? A Personal Reflection. // *Evaluation*, Jan 2004; vol. 10: pp. 7-15
45. Toulemonde, Jacques, „Evaluation culture(s) in Europe: differences and convergence between national practices”, 2000. http://www.danskevalueringsselskab.dk/pdf/Toulemonde_paper.pdf
46. Toulemonde, Jacques and Bjornkilde, Thomas, „Building Evaluation Capacity: Experience and Lessons in Member States and Acceding Countries.”, Budapest, 26/27 June 2003. http://ec.europa.eu/regional_policy/sources/docconf/budapeval/work/toulemonde.doc
47. Toulemonde J., Summa H. and Usher N. „Three layers of quality assurance: would this help provide EU policy makers with the evaluative information they need?”. Paper presented in European Evaluation Society Conference, 10-12 October 2002.
48. Walter I, Nutles S., Davies H., 2005. What works to promote evidence-based practice, *Evidence and Policy* Vol. 1 (3), p.335-63
49. Weiss, C.H. (1999) “The Interface between Evaluation and Public Policy”, *Evaluation* 5(4): 468-86.
50. Weiss, Carol H., *Evaluation* (1998), New Jersey: Prentice-Hall.
51. World Bank (2002) *2002 Annual Report on Evaluation Capacity Development*. Washington, DC: World Bank
52. World Bank „Monitoring and Evaluation: Some Tools, Methods and Approaches“, 2004. [http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/a5efbb5d776b67d285256b1e0079c9a3/\\$FILE/MandE_tools_methods_approaches.pdf](http://lnweb18.worldbank.org/oed/oeddoclib.nsf/24cc3bb1f94ae11c85256808006a0046/a5efbb5d776b67d285256b1e0079c9a3/$FILE/MandE_tools_methods_approaches.pdf)

Web Sources

1. European Commission. EVALSED: the online and interactive resource for the evaluation of socio-economic development. Published on the Internet at: www.evaled.com
2. European Policy Research Centre. <http://www.eprc.strath.ac.uk/iqnet/news.cfm>
3. Policy Hub policy making and delivery across government www.policyhub.gov.uk/
4. Evidence Network (ESRC Centre for Evidence-Based Policy and Practice) <http://www.evidencenetwork.org/>
5. MandE - news service focusing on developments in monitoring and evaluation methods relevant to social development projects and programmes <http://www.mande.co.uk/>
6. Belgium – Wallonie Evaluation Society <http://www.prospeval.org/>
7. Danish Evaluation Society <http://www.danskevalueringsselskab.dk/>
8. Finnish Evaluation Society <http://www.finnishevaluationsociety.net/>
9. French Evaluation Society <http://www.sfe.asso.fr/>
10. German Evaluation Society <http://www.degeval.de/>
11. Hungarian Evaluators Network mate.vincze@hu.pwc.com
12. Italian Evaluation Society <http://www.valutazione.it/>
13. Irish Evaluation Network http://www.policyinstitute.tcd.ie/Irish_Evaluation_Network.html
14. Polish Evaluation Society <http://www.pte.org.pl/>
15. Romanian Evaluation Society <http://www.evalrom.ro>
16. Spanish Public Policy Evaluation Society <http://www.idr.es/>
17. Swedish Evaluation Society <http://www.svuf.nu/>

18. UK Evaluation Society <http://www.evaluation.org.uk/>
19. European Evaluation Society <http://www.europeanevaluation.org/>
20. UNDP Evaluation office <http://www.undp.org/eo/>
21. UN Evaluation Forum <http://www.uneval.org/>
22. World Bank Independent Evaluation Group website - <http://www.worldbank.org/ieg/ecd/>

ANNEX 1. QUESTIONNAIRES FOR BENCHMARKING EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES

QUESTIONNAIRE FOR THE STRUCTURAL AND COHESION FUND ADMINISTRATORS WITH RESPONSIBILITY FOR EVALUATION

QUESTIONNAIRE ON THE STRUCTURAL FUND EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES

You are invited to take part in online survey about Structural and Cohesion Fund evaluation capacity in the EU Member States.

This survey forms part of a research project “*Study on Benchmarking the Development of Evaluation Capacity in the EU Member States*” focusing on the Structural and Cohesion Fund administrations. This project is been carried out by public company ESTEP (www.estep.lt) on behalf of the European Commission.

The purpose of this questionnaire is to gather information about the capacity to evaluate programmes funded by the EU Structural and Cohesion Funds in the Member States. Information from this questionnaire will help us to get a comparative view of the Structural and Cohesion Fund evaluation situation in the Member States and will also provide insights as to how the evaluation capacity can be further developed.

Below are some guidelines for completing the survey:

- The web address for the online questionnaire is: <web link>
- The questionnaire may be accessed from any computer with an Internet connection.
- The survey consists of 41 questions and should take approximately 20 minutes to complete.
- You may hold up filling out the questionnaire by clicking on the “STOP” button. Later when you return to the questionnaire you will continue the survey at the point where you stopped.
- We would be grateful, if you could respond to the questionnaire **by Tuesday, 5 June 2007**.
- If you have any queries, please contact Ms Laura Indriliunaite, e-mail estep@estep.lt or tel. (+370) 5269 0116.

We thank you in advance for your willingness to participate in this survey. Your responses will remain confidential and the results of the survey will be pooled for analysis purposes (such that individual responses cannot be identified). The results of the research will be presented during one of the DG REGIO Evaluation Network meetings and DG EMPL Partnership Meetings.

INFORMATION ABOUT THE RESPONDENT

In this Section, we kindly ask you to give some information about yourself and institution you are representing.

Name

EU Member State

1. Please, indicate the EU Member State in which you work.

Austria	<input type="checkbox"/>
Belgium	<input type="checkbox"/>
Bulgaria	<input type="checkbox"/>
Cyprus	<input type="checkbox"/>
Czech Republic	<input type="checkbox"/>
Denmark	<input type="checkbox"/>
Estonia	<input type="checkbox"/>
Finland	<input type="checkbox"/>
France	<input type="checkbox"/>
Germany	<input type="checkbox"/>
Greece	<input type="checkbox"/>
Hungary	<input type="checkbox"/>
Ireland	<input type="checkbox"/>
Italy	<input type="checkbox"/>
Latvia	<input type="checkbox"/>
Lithuania	<input type="checkbox"/>
Luxembourg	<input type="checkbox"/>
Malta	<input type="checkbox"/>
Netherlands	<input type="checkbox"/>
Poland	<input type="checkbox"/>
Portugal	<input type="checkbox"/>
Romania	<input type="checkbox"/>
Slovakia	<input type="checkbox"/>
Slovenia	<input type="checkbox"/>
Spain	<input type="checkbox"/>
Sweden	<input type="checkbox"/>
United Kingdom	<input type="checkbox"/>

Institution

2. What is the name of institution at which you work?

3. What role does this institution play in relation to the management/implementation of the Structural and Cohesion Funds in your country?

Managing Authority at national level (CSF / SPD / NSRF)	<input type="checkbox"/>	
Operational Programme Managing Authority	<input type="checkbox"/>	→ Please specify the Operational Programme concerned
Intermediate body	<input type="checkbox"/>	→ Please specify the Operational Programme concerned

Other (please, specify):

--

4. What is your position / office held?

Head of the Department / Directorate/ Unit	<input type="checkbox"/>
Senior officer / coordinator	<input type="checkbox"/>

Other (please specify):

--

A. THE DEMAND SIDE

In this Section A we ask questions about how the Structural and Cohesion Fund evaluation function is organised and managed in your country/region and the resources allocated to it. One question concerns the quality of monitoring systems.

ORGANISATION OF EVALUATION FUNCTION

A1. Which directorate, department or unit is responsible for Structural and Cohesion Fund evaluation within the Managing Authority at national level?

Please specify:

--

Do not know	<input type="checkbox"/>
-------------	--------------------------

A2. Is there a separate dedicated unit with responsibility for the Structural and Cohesion Fund evaluation function for the programme (NSRF or Operational Programme) that you are associated with?

No	<input type="checkbox"/>	
Yes	<input type="checkbox"/>	→ <u>please specify</u>
Do not know	<input type="checkbox"/>	

A3. Is there a document (e.g. procedures manual or similar) that defines institutional responsibilities for Structural and Cohesion Fund evaluation function?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
In preparation	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

A4. Are there formal mechanisms for coordination of Structural and Cohesion Fund evaluation at national level?

There are no formal mechanisms	<input type="checkbox"/>
Evaluation Steering Group / Committee	<input type="checkbox"/>
Dedicated institution with responsibility for coordination	<input type="checkbox"/>
Other	→ <u>Please specify</u>
Do not know	<input type="checkbox"/>

EVALUATION RESOURCES

A5. What is the total amount of financial allocations foreseen for the evaluation of the programme (NSRF or Operational Programme) that your institution is associated with?

	2000-2006 / or 2004-2006 period	2007-2013 period
Please specify to nearest €million	specify	specify
Do not know / not applicable	<input type="checkbox"/>	<input type="checkbox"/>

A6. How many persons are allocated to the evaluation function for the programme with which your institution is associated?

Please specify:

--

Do not know	<input type="checkbox"/>
-------------	--------------------------

A7. Do you consider that the financial and human resources devoted to the evaluation of Structural and Cohesion Fund programmes in your country are adequate?

	Inadequate	Less than adequate	Adequate	Do not know
Financial resources	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Human resources	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A8. Is there a supply of training and education options in the field of evaluation for public administration officials in your country?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

A9. Are you aware of any evaluation capacity building activities (e.g. training, conferences, evaluation guides) carried out for public administration officials during the period from 2004 to 2007?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>

→ If yes, please specify the type of capacity building activity [multiple selection]:

Trainings	<input type="checkbox"/>
Workshops/seminars	<input type="checkbox"/>
Development of evaluation guides	<input type="checkbox"/>
Awareness raising activities (e.g. conferences)	<input type="checkbox"/>
Other activities	→ <u>Please specify</u>

A10. Have you participated in any evaluation capacity building activity in the period from 2004 to 2007?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>

→ If yes, please specify the type of capacity building activity [multiple selection]:

Trainings	<input type="checkbox"/>
Workshops, seminars	<input type="checkbox"/>
Development of evaluation guides	<input type="checkbox"/>
Awareness raising activities (e.g. conferences)	<input type="checkbox"/>
Other activities	→ <u>Please specify</u>

A11. How would you assess capacity and skills of public administration officials responsible for Structural and Cohesion Fund evaluation function in the following areas:

	Very poor	Poor	Fair	Good	Very good	Do not know / not applicable
Drafting Terms of Reference for evaluation assignments	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Steering evaluation projects performed by external evaluators	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Conducting internal evaluations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Assessing quality of the evaluation reports	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A12. How many evaluations were undertaken of the programme (NSRF or Operational Programme) with which your institution is associated between 2004 and 2006?

Please specify the number:

--

Do not know	<input type="checkbox"/>
-------------	--------------------------

A13. Did these evaluations include evaluations other than those mandatory according to the EU regulations (*i.e.* ex ante, mid-term and update evaluations)?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

EFFICIENCY OF EVALUATION MANAGEMENT

A14. Has an evaluation plan been drawn up in respect of the programme (NSRF or Operational Programme) with which your institution is associated? [multiple selection]

		No	Yes	Do not know
2000-2006 / or 2004-2006	Evaluation plan at national level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Evaluation plans at the OP level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2007-2013	Evaluation plan at national level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	Evaluation plans at the OP level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A15. Based on your institution's most recent experience in commissioning evaluation, what is the approximate time period from the point at which agreement is reached on Terms of References to the appointment of evaluator?

Please indicate the number of weeks:

--

Do not know	<input type="checkbox"/>
-------------	--------------------------

A16. Are there evaluation steering groups / committees established? (multiple selection)

	No	Yes	Do not know
Evaluation steering group / committee at the national level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Evaluation steering group / committee at the programme level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
For individual evaluation projects	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

A17. Is there an approved set of quality standards against which evaluation reports are assessed?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

QUALITY OF MONITORING SYSTEMS

A18. As regards the programme (NSRF or OP) that your organisation is associated with, how long does it take for the monitoring system to provide monitoring data in respect of the previous year?

Less than 3 months	<input type="checkbox"/>
3-6 months	<input type="checkbox"/>
More than 6 months	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B. THE SUPPLY SIDE

In this Section B we ask about the data used for Structural and Cohesion Fund evaluations and the availability and quality of evaluation services.

AVAILABILITY OF SOCIO-ECONOMIC DATA

B1. Is key official statistical economic and social data (e.g. GDP, employment, R&D investment, etc.) available to evaluators?

	No	Yes	To some extent	Do not know
Data at national level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data at regional (NUTS II) level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

AVAILABILITY AND QUALITY OF EVALUATION EXPERTISE

B2. Do you consider that the supply of Structural and Cohesion Fund evaluation services and expertise is adequate in the national market?

Inadequate	<input type="checkbox"/>
Less than adequate	<input type="checkbox"/>
Adequate	<input type="checkbox"/>
Depends on the sector or type	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B3. Are foreign consulting companies involved in providing evaluation services in the national evaluation market?

Not at all	<input type="checkbox"/>
To a limited extent	<input type="checkbox"/>
To a significant extent	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B4. Are universities / research institutes involved in providing evaluation services?

Not at all	<input type="checkbox"/>
To a limited extent	<input type="checkbox"/>
To a significant extent	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B5. Is there an internal evaluation unit which carries out internal evaluations of the programme that your institution is associated with?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B6. If yes → How many internal evaluations were undertaken between 2004 and 2006?

Please specify the number:

--

Do not know	<input type="checkbox"/>
-------------	--------------------------

B7. In general, how would you rate the quality of evaluation reports carried out by external evaluators over recent years?

Very poor	<input type="checkbox"/>
Poor	<input type="checkbox"/>
Fair	<input type="checkbox"/>
Good	<input type="checkbox"/>
Very good	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B8. How would you rate the thematic expertise of the external suppliers of evaluation services in the following areas?

	Very poor	Poor	Fair	Good	Very good	Do not know
Horizontal principles (sustainable development, regional development, equal opportunities, information society)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Infrastructure	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Human resource development	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Enterprise, competitiveness and innovation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Agriculture, rural development and fisheries	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B9. How would you rate the methodological expertise of the suppliers of evaluation services?

	Very poor	Poor	Fair	Good	Very good	Do not know
Qualitative methods	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

(e.g. interviews, focus groups, desk research, case studies, expert panels, Delphi method)						
Macroeconomic modelling	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quantitative methods (e.g. statistical analysis, cost-benefit analysis, cost-effectiveness analysis)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Experimental/quasi-experimental methods (e.g. control group approaches etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

C. DISSEMINATION AND UTILISATION

In this Section C we examine how Structural and Cohesion Fund evaluation findings are disseminated and utilised in your country/region.

ACCESS TO EVALUATION OUTPUTS

C1. Are Structural and Cohesion Fund evaluation reports publicly available?

Reports are not publicly available	<input type="checkbox"/>
Some reports are publicly available	<input type="checkbox"/>
Most reports are publicly available	<input type="checkbox"/>
All reports are publicly available	<input type="checkbox"/>
Only report summaries are publicly available	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C2. Are Structural and Cohesion Fund evaluation reports placed on the internet?

No	<input type="checkbox"/>
Some	<input type="checkbox"/>
Most	<input type="checkbox"/>
All	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

ADDRESSING EVALUATION RESULTS

C3. Is there an organisation, unit or committee responsible for considering or taking forward the findings and recommendations of completed evaluation reports?

No	<input type="checkbox"/>	
Yes	<input type="checkbox"/>	→ <u>Please specify</u>
Do not know	<input type="checkbox"/>	

C4. Are there procedures for following-up or tracking the implementation of agreed evaluation recommendations?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

IMPACTS ON PROGRAMME DESIGN

C5. How would you rate the influence of ex ante evaluation and its findings as regards the formulation of the national strategy for Structural and Cohesion funding in 2007-2013?

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C6. How would you rate the influence of ex ante evaluation in terms of the allocation of resources to different programmes and priorities for the period of 2007-2013?

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C7. How would you rate the influence of ex ante evaluation on the quantification of objectives and indicator selection for programmes in 2007-2013?

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

D. INSTITUTIONALISATION OF EVALUATION IN WIDER GOVERNANCE SYSTEM

In this Section D we explore a range of more general issues related to wider governance environment in the country and how favourable it is to evaluation practice.

EVALUATION IN STATE BUDGETARY PROCESS

D1. In your opinion, what is the role of evaluation in the national budgetary process?

Not important	<input type="checkbox"/>
Limited importance	<input type="checkbox"/>
Important	<input type="checkbox"/>
Very important	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

EVALUATION IN SECTORAL STRATEGY / POLICYMAKING

D2. To what extent are evaluations used in strategy and policy formulation at sectoral / ministry level?

Never	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Often	<input type="checkbox"/>
Always	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

EXISTENCE OF MANDATORY EVALUATION REQUIREMENTS

D3. Is there a legal basis which sets out mandatory requirements for evaluation of national programmes?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

PUBLIC POLICY RESEARCH BASE

D4. Are universities / research institutes / think-tanks involved in public policy research?

Not at all	<input type="checkbox"/>
To a limited extent	<input type="checkbox"/>
To a significant extent	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

INTERACTION BETWEEN EU STRUCTURAL AND COHESION FUND EVALUATION REQUIREMENTS AND NATIONAL ADMINISTRATIVE CULTURE

D5. What influence have the EU Structural and Cohesion Fund evaluation requirements had on the evaluation practice and capacity in your country?

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

D6. Would you say that the administrative and political culture in your country is favourable to the utilisation of evaluation in the public policymaking process?

Unfavourable	<input type="checkbox"/>
Somewhat unfavourable	<input type="checkbox"/>
Neutral	<input type="checkbox"/>
Broadly favourable	<input type="checkbox"/>
Yes, most conducive	<input type="checkbox"/>

ROLE OF CIVIL SOCIETY AND OTHER PLAYERS

D7. Are the evaluation reports and findings utilised in wider public debates?

Never	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Often	<input type="checkbox"/>
Always	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

Thank you very much for completing this questionnaire.

If you wish to add any comments concerning the questionnaire or Structural and Cohesion Fund evaluation issues, please use the space below.

--

QUESTIONNAIRE FOR THE EVALUATION MARKET PARTICIPANTS

QUESTIONNAIRE ON STRUCTURAL FUND EVALUATION CAPACITY DEVELOPMENT IN THE EU MEMBER STATES

You are invited to take part in online survey about Structural and Cohesion Fund evaluation capacity in the EU Member States. The survey targets those people who have first-hand experience in evaluation of Structural and Cohesion Funds at programme level.

This survey forms part of a research project “*Study on Benchmarking the Development of Evaluation Capacity in the EU Member States*” focusing on the Structural and Cohesion Fund administrations. This project is being carried out by public company ESTEP (www.estep.lt) on behalf of the European Commission.

The purpose of this questionnaire is to gather information about the capacity to evaluate programmes funded by the EU Structural and Cohesion Funds in the Member States. Information from this questionnaire will help us to get a comparative view of the Structural and Cohesion Fund evaluation situation in the Member States and will also provide insights as to how the evaluation capacity can be further developed.

Below are some guidelines for completing the survey:

- The web address for the online questionnaire is: *<web link >*
- The questionnaire may be accessed from any computer with an Internet connection.
- The questionnaire consists of 19 questions and should take approximately 10 minutes to complete
- You may hold up filling out the questionnaire by clicking on the “STOP” button. Later when you return to the questionnaire you will continue the survey at the point where you stopped.
- We would be grateful, if you could respond to the questionnaire **by Tuesday, 5 June 2007**.
- If you have any queries, please contact Ms Laura Indriliunaite, e-mail estep@estep.lt or tel. (+370) 5269 0116.

We thank you in advance for your willingness to participate in this survey. Your responses will remain confidential and the results of the survey will be pooled for analysis purposes (such that individual responses cannot be identified).

INFORMATION ABOUT THE RESPONDENT

In this Section, we kindly ask you to give some very general information about yourself, institution you are representing and your Structural Funds evaluation experience.

Name

--

Organisation

1. What is the name of organisation you represent?

--

2. What is the type of organisation you represent?

Private sector organisation / Consultancy firm	<input type="checkbox"/>
Academic / Research institution	<input type="checkbox"/>
Non-governmental organisation	<input type="checkbox"/>
Other	please, specify

3. Please, indicate your home country.

Austria	<input type="checkbox"/>
Belgium	<input type="checkbox"/>
Bulgaria	<input type="checkbox"/>
Cyprus	<input type="checkbox"/>
Czech Republic	<input type="checkbox"/>
Denmark	<input type="checkbox"/>
Estonia	<input type="checkbox"/>
Finland	<input type="checkbox"/>
France	<input type="checkbox"/>
Germany	<input type="checkbox"/>
Greece	<input type="checkbox"/>
Hungary	<input type="checkbox"/>
Ireland	<input type="checkbox"/>
Italy	<input type="checkbox"/>
Latvia	<input type="checkbox"/>
Lithuania	<input type="checkbox"/>
Luxembourg	<input type="checkbox"/>
Malta	<input type="checkbox"/>
Netherlands	<input type="checkbox"/>
Poland	<input type="checkbox"/>
Portugal	<input type="checkbox"/>
Romania	<input type="checkbox"/>

Slovakia	<input type="checkbox"/>
Slovenia	<input type="checkbox"/>
Spain	<input type="checkbox"/>
Sweden	<input type="checkbox"/>
United Kingdom	<input type="checkbox"/>

Structural and Cohesion Fund evaluation experience

4. What is your most recent Structural and Cohesion Funds evaluation project that you have carried out in your home country?

Please specify the title of the Structural and Cohesion Fund evaluation project:

5. Please, indicate the type of your most recent Structural and Cohesion Fund evaluation that you have carried out in your home country:

	Objective 1	Objective 2	ESF specific
Ex ante evaluation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mid-term (including update) evaluation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thematic evaluation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ex post evaluation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

6. Please specify the year the evaluation project was completed:

The questionnaire below includes both general questions on evaluation practice in your home country as well as more specific questions relating to the most recent evaluation experience in your home country.

We kindly ask you to reply to the following questions with regard to the country which you have indicated above as your home country and your most recent evaluation project carried out in this country.

A. THE DEMAND SIDE

In this Section A we ask a few questions about how the Structural and Cohesion Funds evaluation function is organised and managed in your country/region and the resources allocated to it. One question concerns the monitoring systems.

ORGANISATION OF EVALUATION FUNCTION

A1. In your view, how clearly are the institutional responsibilities and procedures for Structural and Cohesion Fund evaluation function defined in your country?

Not defined	<input type="checkbox"/>
Poorly-defined or unclear	<input type="checkbox"/>
Adequately-defined	<input type="checkbox"/>
Well-defined	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

Comments

--

EVALUATION RESOURCES

A2. Based on your most recent evaluation experience, how would you assess the skills of public administration officials responsible for Structural and Cohesion Fund evaluation function in the following areas?

	Very poor	Poor	Fair	Good	Very good	Do not know / not applicable
Drafting Terms of Reference for evaluation assignments	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Steering evaluation projects performed by external evaluators	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

QUALITY OF MONITORING SYSTEMS

A3. Was the necessary monitoring data available for your last Structural and Cohesion Fund evaluation?

No, monitoring data was not available	<input type="checkbox"/>
Some monitoring data was available	<input type="checkbox"/>
Most monitoring data was available	<input type="checkbox"/>
Yes, all the necessary monitoring data was available	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

B. THE SUPPLY SIDE

In this Section B we ask about the availability of data used for Structural and Cohesion Fund evaluations and supply of evaluation training and education options on the domestic market.

AVAILABILITY OF SOCIO-ECONOMIC DATA

B1. Was necessary official statistical economic and social data (e.g. GDP, employment, R&D investment, etc.) available for your last Structural and Cohesion Fund evaluation?

	No	Yes	To some extent	Do not know
Data available at national level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Data available at regional (NUTS II) level	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

TRAINING AND EDUCATION IN THE FIELD OF EVALUATION

B2. Is there a supply of training and education options in the field of evaluation in your country?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C. DISSEMINATION AND UTILISATION

In this Section C we examine how Structural and Cohesion Fund evaluation findings are disseminated and utilised in your country.

ACCESS TO EVALUATION OUTPUTS

C1. Are Structural and Cohesion Fund evaluation reports publicly available?

Reports are not publicly available	<input type="checkbox"/>
Some reports are publicly available	<input type="checkbox"/>
Most reports are publicly available	<input type="checkbox"/>
All reports are publicly available	<input type="checkbox"/>
Only report summaries are publicly available	<input type="checkbox"/>

C2. Are Structural and Cohesion Fund evaluation reports placed on the internet?

No	<input type="checkbox"/>
Some	<input type="checkbox"/>

Most	<input type="checkbox"/>
All	<input type="checkbox"/>

UTILISATION OF EVALUATION FINDINGS

C3. Were the findings and recommendations of your most recent evaluation utilised?

None	<input type="checkbox"/>
Some	<input type="checkbox"/>
Most	<input type="checkbox"/>
All	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

IMPACTS ON PROGRAMME DESIGN

C4. Have you carried out an ex ante evaluation for Structural and Cohesion funding in 2007-2013?

No	<input type="checkbox"/>
Yes	<input type="checkbox"/>

C5. How would you rate the influence of ex ante evaluation as regards the formulation of the national strategy for Structural and Cohesion funding in 2007-2013? [asked only if responded “yes” to question C4]

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C6. How would you rate the influence of ex ante evaluation in terms of resources to different priorities and programmes for the period of 2007-2013? [asked only if responded “yes” to question C4]

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

C7. How would you rate the influence of ex ante evaluation for the quantification of objectives and indicator selection for programmes in 2007-2013? [asked only if responded “yes” to question C4]

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>

Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

D. INSTITUTIONALISATION OF EVALUATION IN WIDER GOVERNANCE SYSTEM

In this Section D we explore a range of more general issues related to wider governance environment in the country and how favourable it is to evaluation practice.

EVALUATION IN STATE BUDGETARY PROCESS

D1. In your opinion, what is the role of evaluation in the national budgetary process?

Not important	<input type="checkbox"/>
Limited importance	<input type="checkbox"/>
Important	<input type="checkbox"/>
Very important	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

EVALUATION IN SECTORAL STRATEGY / POLICYMAKING

D2. To what extent are evaluations used in strategy and policy formulation at sectoral / ministry level?

Never	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Often	<input type="checkbox"/>
Always	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

EVALUATION SOCIETY OR NETWORK

D3. Are you a member of a national professional body (association, society) of evaluators?

Yes	<input type="checkbox"/>	→ Please specify the <u>country</u>
No	<input type="checkbox"/>	

D4. In how many evaluation events (conferences, seminars, etc.) did you participate in the period from 2004 to 2007?

Evaluation events in your country	Please indicate the number
Evaluation events in other EU Member States	Please indicate the number

INTERACTION BETWEEN EU STRUCTURAL AND COHESION FUND EVALUATION REQUIREMENTS AND NATIONAL ADMINISTRATIVE CULTURE

D5. In your view, what influence have the EU Structural and Cohesion Funds evaluation requirements had on the evaluation practice and capacity in your country?

No influence	<input type="checkbox"/>
Minor influence	<input type="checkbox"/>
Moderate influence	<input type="checkbox"/>
Strong influence	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

D6. Would you say that the administrative and political culture in your country is favourable to the utilisation of evaluation in the public policymaking process?

Unfavourable	<input type="checkbox"/>
Somewhat unfavourable	<input type="checkbox"/>
Neutral	<input type="checkbox"/>
Broadly favourable	<input type="checkbox"/>
Yes, most conducive	<input type="checkbox"/>

ROLE OF CIVIL SOCIETY AND OTHER PLAYERS

D7. Are the evaluation reports utilised in public debates?

Never	<input type="checkbox"/>
Sometimes	<input type="checkbox"/>
Often	<input type="checkbox"/>
Always	<input type="checkbox"/>
Do not know	<input type="checkbox"/>

Thank you very much for completing this questionnaire.

If you wish to add any comments concerning the questionnaire or Structural and Cohesion Fund evaluation issues, please use the space below.

ANNEX 2. CASE STUDY REPORTS

Report on England Case Study

(1) Introduction and Methodology

This report presents the results of the case study of evaluation capacity in England. The case study is based on a review of relevant documentation and a series of semi-structured face-to-face and telephone interviews carried out over the week 25 to 29th June 2007. The full list of documents reviewed is set out in the annex; in the main the documents consulted were reviews of various aspects of the state of evaluation practice in the UK as a whole. A total of 5 persons were interviewed for the purposes of the case study. A joint face-to-face interview was conducted with representatives of (the Department for) Communities and Local Government and the then Department of Trade and Industry (now Department for Business, Enterprise and Regulatory Reform) in London on 25th June. At the time of writing (August 2007), the Department for Business, Enterprise and Regulatory Reform (DBERR) has overall policy responsibility for the Structural and Cohesion Funds in the UK as a whole whereas Communities and Local Government (CLG) acts as managing authority for ERDF-funded programmes in England (the transfer of the CLG functions to DBERR is currently under discussion). Telephone interviews were conducted with a representative of the ESF Evaluation Team at the Department for Work and Pensions (based in Sheffield) on 26th June (the team responsible for the evaluation of Objective 3 ESF funding in the 2000 to 2006 period and for the ESF Operational Programme for England for 2007 to 2013). The final (telephone) interview was conducted with Mr. Jack Malan, Centre for Strategy and Evaluation Services on 28th June who has significant experience in undertaking evaluations of EU-funded programmes in England.

The documentary analysis and the interviews provided a wealth of useful information on the various components (i.e., dimensions, criteria, sub-criteria and indicators) included in the evaluation capacity development benchmarking framework. In some cases, hard quantitative data was not readily to hand (e.g., indicators relating to evaluation budgets, procurement intervals and the number of firms active in the evaluation market) although some general impressions or views on the issues concerned were often conveyed.

(2) Background Information: Structural Funds in England

For the 2007 to 2013 programming period, most of the English regions are eligible under the Regional Competitiveness and Employment objective. Two regions qualify for phasing-in support under this objective (Merseyside and South Yorkshire, both objective 1 regions in the 2000 to 2006 period) while Cornwall and Isles of Scilly is the only region in England eligible under the Convergence objective (also an objective 1 region in 2000 to 2006). There will be separate ERDF programmes for each of the ten regions (9 under the competitiveness objective and 1 convergence programme for Cornwall and Isles of Scilly). As mentioned above, the Managing Authority for these programmes will be Communities and Local Government (CLG) with day-to-day management and delivery of the programmes delegated to the Regional Development Agencies (RDAs). For the ESF, a single operational programme with ESF funding of €3090 million will cover England as a whole, with managing authority responsibility resting with the Department for Work and Pensions (DWP).

(3) Demand Side

The analysis is presented below under the 4 criteria included under the demand-side dimension of the evaluation capacity development benchmarking framework as follows:

- The integration, location and organisation of the Structural Funds (SF) evaluation function;
- Evaluation resources;
- Efficiency of evaluation management;
- Quality of monitoring systems.

Integration, Location and Organisation of Evaluation Function

As mentioned above, overall policy responsibility for the SF in the UK rests with the DBERR. In terms of EU evaluation matters, the DBERR participates in the work of the DG Regional Policy Evaluation Network but in the main responsibility for meeting SF evaluation requirements in England lies with the ERDF and ESF managing authorities (Communities and Local Government and the Department of Work and Pensions respectively).³⁹ In the case of the ERDF programmes, CLG coordinates the evaluation process and is represented on evaluation steering committees that operate at the level of each regional OP and RDA (or regional government offices in the 2000-2006 period). The programme secretariats attached to the RDAs/local government offices are then responsible for managing individual programme-level evaluation projects (generally the evaluation and monitoring functions are closely linked at this level) although CLG has tended to play a role in centrally organising some aspects of the procurement process. Thus, responsibility for evaluation of ERDF programmes is to some extent split between CLG and the management entities at regional level.

As mentioned earlier, managing authority responsibility for the single ESF programme in England for 2007-2013 (and for the 2000-2006 Objective 3 programme) lies with the Department for Work and Pensions (DWP). Responsibility for evaluation of these programmes rests with a dedicated ESF Evaluation Team based in the DWP but functionally independent of the Managing Authority. The Team participates in the work of the UK ESF Evaluation Standing Group which oversees evaluation of all ESF programmes in the UK. This committee includes representatives from England, Scotland, Wales, Northern Ireland, the European Commission and the ESF Division (Managing Authority) in the DWP. The aim of the group is to promote consistency and quality in the way that ESF programmes are evaluated and to share good practice in evaluation methodology. For the 2007 to 2013 ESF programme for England, an evaluation sub-committee of the Programme Monitoring Committee will consider the evaluation strategy and receive reports from the ESF Evaluation Team on the progress and results of evaluation activity. As regards monitoring, the Team's role is one of analysing monitoring data and feeding this into the work of the Monitoring Committee for the England OP (the monitoring data is compiled by a separate group of DWP statisticians). Overall, therefore, the monitoring and evaluation functions are quite closely linked.

Evaluation Resources

In terms of financial resources, there are no separate, ring-fenced OP evaluation budgets as such and data on evaluation costs for the 2000-2006 period were not readily to hand. The average cost of the mid-term evaluations for ERDF programmes was estimated at about £100,000 (equivalent to about €150,000). For the ESF programme, follow-up beneficiary surveys cost around £500,000 (about €750,000).

For ERDF programmes – which are individually quite small in terms of financial weight - there are no dedicated evaluation units or officials assigned exclusively to commissioning or managing the evaluation process. The ESF Evaluation Team currently comprises three social research staff, equating to two full-time equivalent staff (one full-time and two part-time). The intention is to

³⁹ It is important to note that the DBERR also plays an important role in promoting evaluation of regional policy interventions in England, particularly the activities of the Regional Development Agencies, who are involved in the implementation of EU programmes in England. The DBERR has published evaluation guidance for the RDAs (see DBERR occasional paper no. 2: *Evaluating the impact of England's Regional Development Agencies: Developing a Methodology*).

expand the staff complement in the near future to include an additional social researcher, statistician and economist. All of the current staff are trained researchers and are members of Government Social Research (GSR) whose role is discussed below.

In terms of the volume of evaluation outputs, the mid-term and update evaluations were completed for all of the ERDF 2000-2006 programmes (In the case of the update evaluation, a composite evaluation covering all programmes was compiled by CLG based on inputs from the OP level.) Over and above these requirements of the Council Regulations, a thematic evaluation was carried out based on case studies across different (ESF and ERDF) programmes (commissioned by CLG). The main focus of this study was on approaches and interventions that worked well with the aim of providing a fuller evidence base to inform the development of future regional economic policies and programmes.⁴⁰ Ex ante evaluations have also been carried out for each of the 2007-2013 ERDF programmes.

For the ESF programmes in England, a substantial volume of evaluation and research work – going well beyond the basic regulatory requirements – is carried out. As well as the obligatory mid-term and update evaluations for the 2000-2006 period, follow-surveys of ESF beneficiaries (both individuals and companies) are undertaken regularly as well as other research and ad-hoc evaluations. For example, in 2005, some 4 evaluation/research outputs were commissioned comprising an evaluation of an Objective 3 Global Grants Programme, a follow-up survey of ESF beneficiaries, a profile of inactive ESF beneficiaries and research into multiple disadvantaged groups.

Efficiency of Evaluation Management

For ESF programmes, the DWP-based ESF Evaluation Team is responsible for producing an ongoing evaluation strategy and plan. For the 2007 to 2013 period, a draft plan has already been produced covering both the convergence and competitiveness objectives. (The ESF OP as submitted to the Commission provides an outline of the main issues that will be the focus of the Plan.) This will be considered in due course by the evaluation sub-committee of the England ESF Programme Monitoring Committee.

For the ERDF programmes, matters are not as advanced but it should be noted that the average OP size is much smaller than is the case for the national ESF OP. A standard text in each of the draft 2007-2013 OPs indicates that “the Managing Authority will undertake such additional evaluations as deemed necessary”. Separately, the DBERR is pressing the RDAs (who will manage the ERDF programmes) to undertake evaluations for the purposes of the Treasury-led Comprehensive Spending Review process (this is discussed further below under the governance dimension).

Terms of reference for individual evaluation projects are developed by the ESF Evaluation Team for ESF activity. For the ERDF programmes, CLG prepared core terms of reference for the 2000-2006 mid-term and update evaluations and the 2007-2013 ex ante evaluations, drawing on the relevant Commission guidance documents. Generally speaking, the TOR are regarded as being of good quality by external evaluators. No major difficulties with the procurement process (for example in terms of delays in appointing evaluators) were observed; these processes are generally handled centrally at CLG and DWP level.

For both ESF- and ERDF-funded programmes, steering committees are established to oversee individual evaluation projects. Typically, the membership for ERDF evaluations comprises the local government office, RDA, the CLG and some sectoral representatives (e.g., from the voluntary sector) with a total of about 6 members. For the ESF, the Commission (DG Employment) participates occasionally as well as DWP representatives and other stakeholders. For ERDF programmes, quality assessment is carried out on the basis of an adapted version of the

⁴⁰ *Good Practice Guide for English ERDF and ESF Programmes 2007 – 2013, Case Study Evaluations of Wider Impacts of English Objective 1 and 2 Activities*, Final Report by Regeneris Consulting, June 2006

Commission's MEANS quality assessment framework. ESF evaluations are assessed by the steering committee and the ESF Evaluation Team. This assessment draws on the research and evaluation expertise of the Team rather than being based on pre-set quality criteria.

Quality of Monitoring Systems

There are divergent views on the quality of OP monitoring systems. The DWP and CLG provide advice and assistance to programme managers on indicators and have developed standard lists of core performance indicators so as to ensure an element of consistency and stability in the data collected at regional and project level. As noted earlier, the ESF Evaluation Team commissions regular follow-up surveys of beneficiaries which provide a wealth of useful information for monitoring and evaluation purposes.

(4) Supply Side

The analysis below of the supply-side is presented under the two criteria included under the supply dimension of the evaluation capacity development benchmarking framework, namely:

- Quality of socio-economic data resources; and
- Availability and quality of evaluation expertise.

Quality of Socio-Economic Data Resources

In general, the quality and timeliness of the type of contextual, socio-economic indicator data (e.g., GDP, employment, unemployment, R&D investment) that is typically drawn on in evaluations of EU programmes is felt to be of a high standard in England (and for the UK as a whole). Primary responsibility lies with the Office for National Statistics and a wide range of NUTS II regional level data is available from the ONS website (<http://www.statistics.gov.uk>). In addition, the DBERR publishes on an annual basis the *Regional Competitiveness and the State of the Regions* report which includes a wide range of competitiveness indicators for the English regions (data for most indicators is available up to and including 2005 in the recently-published May 2007 report).

Availability and Quality of Evaluation Expertise

The evaluation market in England (and the UK as a whole) is regarded as a competitive one with a considerable number of consultancy firms and other research organisations active in the evaluation sphere. However, the market for evaluation of EU programmes tends to be dominated by a small number of consultancy firms, variously referred to as the "usual suspects" or "usual players". There is limited involvement of the universities. Within central Government, the only internal evaluation unit is the DWP-based ESF Evaluation Team, (separate to the Managing Authority see above) although Government departments generally have a strong internal evaluation and research capacity (discussed below). At regional level (i.e., the local Government Offices and the RDAs), internal evaluation capacity is felt to more limited although the situation varies as between the regions. All informants felt that the quality of evaluation work was generally of a high standard and some observed that this view was one shared by the European Commission.

As regards the availability of education and training options in evaluation, a number of institutions offer courses with an evaluation component to postgraduate level. For researchers working in the government sector, an MSc in Policy Analysis and Evaluation is run in collaboration between the Government Social Research Unit in the Treasury and a college of the University of London. Shorter-duration professional development workshops and training events are organised under the aegis of the UK Evaluation Society (discussed further below).

(5) Dissemination and Utilisation

Evaluation reports are readily available to interested parties with most evaluations available on departmental or local government office websites. For example, all ESF evaluations can be accessed at the England ESF website (<http://www.esf.gov.uk>).

In general, completed evaluation reports are submitted to the relevant programme monitoring committee. There are no formal procedures laid down as to how evaluation reports and recommendations should be dealt with (and tracked where agreed) although there is a general expectation that implementing bodies will take the recommendations seriously (and would be expected to explain why a recommendation cannot be taken forward).

There are some indications of evaluations influencing programme design and implementation. In the case of the ERDF programmes in 2000 to 2006, it was felt that the mid-term evaluations contributed to re-programming in terms of financial re-allocations between OP priorities and the revision of targets. However, some of these financial changes were also driven by N+2 de-commitment concerns. The thematic evaluation carried out by Regeneris (referred to above) was felt to have represented a useful input to the preparation of the 2007-2013 suite of ERDF programmes. In the case of ESF, some of the evaluations carried out in the 2000 to 2006 period (and the ex ante evaluation of the 2007-2013 OP) were felt to have led to a sharper focus in the proposed 2007 to 2013 programmes on the needs of disadvantaged groups such as those not in education, employment and training (the so-called NEETs). The introduction of proposed ESF community grants in the new programme was also influenced by an evaluation of the Objective 3 Global Grants Programme in the 2000 – 2006 period.

(6) Degree of Institutionalisation of Evaluation in the Wider Governance System

This section considers the various criteria specified under the institutionalisation dimension of the evaluation capacity development benchmarking framework as follows:

- The role of evaluation in the national budgetary process;
- Role of evaluation in the formulation of sectoral strategy and policy making;
- The existence of mandatory evaluation requirements;
- The public policy research base;
- National evaluation society;
- Influence of EU evaluation requirements; and
- The role of civil society.

Given that these indicators essentially relate to the national, Member State level, they are considered below at the level of the UK as a whole.

Relative to other EU countries, evaluation and related activities such as social research and performance management play a significant role in the business of government in the UK. In the words of one commentator: “The United Kingdom Government uses a wide range of evaluation methods to ensure that policies, programmes and public services are planned and delivered as effectively and efficiently as possible (Davies, 2004).” In particular, evaluation has become more firmly established since the Labour Government came to power in 1997 with a strong commitment to the practice of evidence-based policy (EBP) making. Since then a series of White Papers and other government publications have stressed the importance of basing policy on sound evidence, good analysis and proper evaluation. These include the White Paper *Modernising Government* (1999) and other Cabinet Office publications including *Professional Policymaking for the 21st century* (1999) and *Better Policy Making* (2001).

Of particular significance has been the introduction of a performance management system for the allocation of budgetary resources. Since 1997 public expenditure has been allocated on a multi-annual basis through the Comprehensive Spending Review (CSR) process. Under this process, each

Government Department negotiates a Public Service Agreement (PSA) every couple of years with the Treasury. The PSAs include a series of objectives and outcome-based performance targets that are linked to an agreed allocation of resources. Progress against the PSAs are then the subject of regular monitoring and evaluation and departments are expected to produce evidence as to the success or otherwise of the PSAs. As one observer has remarked “..the Treasury has been demanding more and more in the way of research evidence to justify a Department’s spending programmes” (Wolf 2006).

The CSR process is complemented by the Treasury’s *Green Book* on economic appraisal and evaluation which applies to policies, programmes and projects. The most recent version was issued by the Treasury in 2003. In its own words, it “constitutes binding guidance for departments and executive agencies” (HM Treasury 2003). The *Green Book* is complemented by the *Magenta Book* (prepared by the Government Social Research Unit, also based at the Treasury) which takes the form of a series of guidance notes on policy evaluation and analysis. More detailed guidance for particular policy areas has also been issued by other government departments, including the DBERR (the evaluation guidance for the RDAs noted earlier), the DWP and the Department of Transport.

These developments have led to a significant increase in the volume of evaluation and research carried out by central government. Martin (2005) notes that the number of social researchers working in government more than doubled between 1997 and 2003 (some 1000 social researchers now work in government)⁴¹ and that the Office of the Deputy Prime Minister (now CLG) increased its budget for research on local and regional government threefold between 1997/8 and 2002/3. A number of major evaluations of policy initiatives in the labour market and welfare policy area have been undertaken over recent years, often using sophisticated experimental designs.

In terms of the governance criteria incorporated in our benchmarking framework, it appears from the above discussion that evaluation (and related practices) play an important part in the budgetary process and in sectoral policy making. The current emphasis attached to evaluation and research in central government reflects the impact of significant reforms in the policy making and resource allocation processes. As noted above, the effect of the *Green Book* is to impose evaluation requirements on government departments and agencies even if technically these provisions do not have the “force of law”. Overall, as concluded by Davies (2004): “The UK Government has a well developed system of policy evaluation ...”⁴²

Focusing on the question of the existence of a public policy research base, the increase in the numbers of social researchers employed in central government was highlighted above. The role of Economic and Social Research Council should also be noted; this is an independent organisation with an annual budget of about £100 million (mostly financed by Government) which funds research and training in economic and social issues. Outside the government sector, a number of think-tanks or research bodies are active in producing policy advocacy papers and research outputs. These include bodies such as DEMOS, the Institute for Public Policy Research and the Joseph Rowntree Foundation.

The UK Evaluation Society (UKES) was established in 1994 and aims “to promote and improve the theory, practice understanding and utilisation of evaluation and its contribution to public knowledge”. According to its website, it has a membership of about 150 individual and corporate members, including evaluation practitioners and commissioners from national and local government, the voluntary sector and other fields. The Society is an active one, holding an annual

⁴¹ See interview with George Clark, retiring Head of Social Research at the DWP, on the Government Social Research website www.gsr.gov.uk

⁴² It should be noted that some commentators (e.g., Martin 2005) have taken a more critical view of recent developments in the UK, arguing for example, that the reforms have given rise to a focus on narrow, target-driven performance measures and an associated inspection culture. One informant referred to an emerging “fatigue” with target-driven approaches.

conference and also organising training events and a programme of regional workshops. The Society has produced good practice guidelines to support the work of evaluators, commissioners of evaluation, evaluation participants and those involved in self-evaluation. It also publishes *The Evaluator* newsletter three times a year.

The findings of the substantial volume of evaluation and research outputs produced in the government sector and by outside think-tanks tend to feed into the media and public debates around government policy. Research and evaluation findings are frequently used for advocacy purposes by civil society actors.

The influence of EU SCF requirements on UK evaluation practice has been very limited, reflecting the reality that the UK had a strong prior evaluation tradition. Evaluation practice in the UK developed independently of EU requirements. When asked about this issue, one informant took the view that, if anything, UK thinking has influenced evaluation practice in other EU countries (mainly the new Member States) and the European Commission's approach to evaluation matters.

(7) Conclusions

The main conclusions from the above analysis are as follows:

- Evaluation and other related practices, including social research and performance measurement, represent an important element of the policy making process in the UK.
- The evaluation of EU programmes benefits from this strong evaluation culture and capacity in central government. On the demand-side, the evaluation of EU programmes in England appears to be well-organised and adequately resourced. The evaluation of ESF programmes is comprehensive in scope and appears to be based on a well-thought out evaluation strategy. ERDF evaluation would appear to be managed in a competent and efficient manner.
- Supply-side factors do not appear to act as a constraint on evaluation of EU programmes in England. A limited number of specialised consultancy firms appear to dominate the market for evaluation of EU programmes. Evaluation outputs appear to be of a good quality, reflecting the expertise of these practitioners.
- Evaluation reports are widely disseminated and appear to have had some influence on the design and implementation of EU programmes.

(8) Key Learning Points

With regard to evaluation capacity benchmarking framework

- Generally speaking, the application of the evaluation capacity benchmarking framework in the case of England did not encounter any serious problems aside from difficulty noted earlier in sourcing quantitative data on some indicators. In overall terms, the framework provided a useful basis on which to organise the fieldwork and to structure this case study report.

From English experience

- There are a number of features of the English (and wider UK) situation that may be of interest to other countries. Of course, the most notable feature of evaluation practice in the UK is its strong integration into the policy cycle and the framework for planning and reporting on public spending. While the focus of this study is on the narrower field of evaluation of Structural and Cohesion Funds, the UK experience will be of interest to other Member States who wish to develop the role of evaluation in the sectoral policy-making and resource allocation spheres.
- In the field of evaluation of the Structural Funds in England, the role of the ESF Evaluation Team based at the DWP is an interesting one. The model of an internal evaluation unit that is attached to - but functionally independent of - the managing authority is one that might be of

interest to other Member States, particularly the newer Member States as they establish evaluation structures for the 2007 to 2013 period (a similar model was also successfully applied in Ireland in previous programming periods).

Bibliography

1. Policy and Programme Evaluation in the United Kingdom: A Reflective State?, in *International Atlas of Evaluation*, Furubo, Rist and Sandahl, 2002
2. Appraisal and Evaluation in Central Government (*The Green Book*), HM Treasury, January 2003
3. *Policy Evaluation in the United Kingdom*, Philip Davies, May 2004
4. Evaluation, Inspection and the Improvement Agenda, Contrasting Fortunes in an Era of Evidence-Based Policy-Making, Steve Martin, *Evaluation* 11(4): 496-504
5. *Assessing the Evidence: The Evaluation of Regional Policy in Europe*, Laura Polverari and John Bachtler, European Policies Research Centre, September 2004

Report on Hungary Case Study

(1) Introduction and Methodology

This report presents results of the case study of Structural Fund evaluation capacity in Hungary. The case study is based on a review of relevant documentation⁴³ and semi-structured face-to-face interviews carried out in Budapest on 12 and 13 June 2007. The full list of documents reviewed is set out in the bibliography. A total of 4 persons were interviewed for the purpose of the case study. Face-to-face interviews were conducted with representatives of the National Development Agency and with the Managing Authority of Human Resource Development Operational Programme (HRDOP). Two interviews were conducted with representatives of private sector companies with experience in Structural Fund evaluation. The results of the evaluation capacity benchmarking survey were also taken into consideration.⁴⁴ The documentary analysis, interviews and survey provided useful information on the various components included in the evaluation capacity development benchmarking framework, structured around four core dimensions as follows:

- Demand, which refers to the commissioning side of the evaluation process;
- Supply;
- Dissemination and utilisation of evaluation outputs;
- Institutionalisation dimension which aims at capturing the wider administrative culture (i.e., outside the Structural Fund administration).

In the report below we firstly present some general information on the Structural Funds in Hungary. This is then followed by the analysis of evaluation capacity.

(2) Background Information: Structural Funds in Hungary

Hungary joined the European Union on 1 May 2004 and became eligible for support from the EU Structural Funds. During the period of 2004-2006 Hungary was eligible for Structural Fund support under Objective 1 and a National Development Plan (NDP) made of 5 sectoral operational programmes (OPs) was implemented.⁴⁵ The Community Support Framework (CSF) Managing Authority (MA) was located in the Office of the Prime Minister, within the Office for the National Development and EU Support (NDO).⁴⁶ The MAs of the different OPs were based at the relevant ministries.

In 2007-2013 six out of the seven Hungarian NUTS II regions fall under the Convergence objective and one (the Central-Hungary region) under the phasing-in Regional Competitiveness and Employment objective.⁴⁷ For the 2007-2013 programming period, Hungary has been allocated an amount of 25.3 billion euros.⁴⁸ Fifteen operational programmes derive from the objectives and priorities of the National Strategic Reference Framework (NSRF), or the so-called New Hungary Development Plan (NHDP). Two OPs are co-financed from the European Social Fund (ESF) and

⁴³ Which was available in English.

⁴⁴ In total 5 representatives from administration and 7 market participants completed the on-line survey questionnaire.

⁴⁵ These OPs were as follows: The Economic Competitiveness Operational Programme (ECOP), Human Resource Development Operational Programme (HRDOP), Environmental Protection and Infrastructure Operational Programme (EPIOP), Agricultural and Rural Development Operational Programme (ARDOP) and The Regional Development Operational Programme (RDOP)

⁴⁶ NDO (under different names) has been in charge for EU funding since 1997. Between 1997 and 2002 the pre-accession programmes were managed by the EU Assistance Coordination Secretariat in the Prime Minister's Office. In 2002 Secretariat for Assistance Coordination was incorporated into newly established Office for National Development Plan and EU Support of Hungary (NDO).

⁴⁷ Hungary. Cohesion Policy 2007-13. // http://ec.europa.eu/regional_policy/atlas2007/fiche/hu_en.pdf

⁴⁸ The Hungarian contribution will reach an amount of € 4.4 billion euros.

thirteen from the European Regional Development Fund (ERDF) and the Cohesion Fund (CF). The fifteen operational programmes comprise seven regional⁴⁹ and eight sectoral programmes⁵⁰.

Significant reform of the Structural Fund administration system was implemented in 2006 with the perspective of 2007-2013 programming period. In July 2006 the new Government established National Development Agency (NDA),⁵¹ which is the legal successor to the NDO.⁵² NDA is responsible for coordination of planning, programming, implementation and evaluation of the NSRF. With the exception of the Agricultural and Rural Development Operational Programme, the NDA has taken over the tasks from all managing authorities of the various OPs.⁵³ As of July 2006 managing authorities of all OPs operate as independent departments of the NDA. The concentration of the Structural Fund management functions at the NDA is seen as a means to improve coordination, facilitate exchange of experience and contribute to transparency and accountability.⁵⁴

(3) Demand Side

The analysis is presented below under the 4 criteria included under the demand-side dimension of the Evaluation Capacity Development Benchmarking Framework as follows:

- The integration, location and organisation of the Structural Funds evaluation function;
- Evaluation resources;
- Efficiency of evaluation management, and
- Quality of monitoring systems.

Integration, Location and Organisation of Evaluation Function

In Hungary the responsibility for organising the evaluation function is shared between the NSRF and OP level. At the time of writing, a document which will set out in detail institutional responsibilities for Structural Fund evaluation function was in the course of preparation. However, a standard text is included in the OPs specifying the overall evaluation arrangements and key functions of the two levels (i.e. the NSRF and OP).

At national level, the most important role with regard to evaluation of Structural Fund programmes is performed by the Department of Evaluation and Methodology under the NDA (the Department).⁵⁵ In the 2007-2013 programming period the Department is responsible for organisation of all the Structural Fund evaluation work and overall coordination of evaluation across the NSRF. The Department provides resources for the implementation of the national evaluation plan and will commissions planned strategic evaluations as well as cross-cutting evaluations that concern more than one OP. It has to ensure methodological consistence of NSRF-level evaluations. In addition, the Department is responsible for organisational development (i.e. training opportunities, development of curriculum, and dissemination of good practice), preparation

⁴⁹ *Convergence objective*: West Pannon Operational Programme; Central Transdanubia Operational Programme; North Hungary Operational Programme; North Great Plain Operational Programme; South Great Plain Operational Programme; South Transdanubia Operational Programme.

Regional competitiveness and employment objective: Central Hungary Operational Programme.

⁵⁰ *Sectoral OPs*: Economic Development OP, Transport OP, Social Renewal OP, Social Infrastructure OP, Environment and Energy OP, State Reform OP, Electronic Public Administration OP, Implementation OP.

⁵¹ The National Development Agency was established by a government decree 130/2006. (VI. 15.) on 1 July, 2006.

⁵² The NDA is headed by the president and vice-presidents. Minister of Local Government and Regional Development supervises the NDA.

⁵³ About National Development Agency.
<http://cordis.europa.eu/erawatch/index.cfm?fuseaction=org.document&uuiid=7D87D84D-DB9A-B9CF-49E2A7A02BFBAF4D>

⁵⁴ As stated in the NHDP, these changes were implemented taking into account the experience of the 2004-2006 period. The deficiencies experienced in the course of implementation of the first NDP included deficient communication between the various levels of the institutional system. // New Hungary Development Plan - P.54.

⁵⁵ In 2004-2006 - the Department for Analyzing, Evaluation and Modelling of the NDO.

of evaluation guidelines and organisation of evaluation events, including those aimed at capacity building.

The OP MAs are responsible for the evaluation of their respective OP and related action plans and are required to maintain adequate capacity for the organisation of evaluations with a performance or operational focus. The MA has to ensure that evaluations of the OP are carried out in accordance with the Community's legislation and that evaluation results are incorporated into the planning and implementation. Each MA makes proposals for the NSRF evaluation plan and is responsible for the implementation of those elements of the plan which concern its respective programme. It drafts the OP evaluation plan and launches, where necessary, evaluations not included in the OP evaluation plan.

In summary, the MAs of the OPs have important sectoral expertise and are responsible for evaluations of their respective OPs. The Department of Evaluation and Methodology under NDA ensures overall coordination of the Structural Fund evaluation activity and ensures a systematic methodological approach. Apart from the Department, coordination is also ensured through the Evaluation Committee. In 2004-2006 the Committee consisted of representatives from the NDO, MAs and ministries

Evaluation Resources

In addition to the evaluation of pre-accession support, evaluations of Structural Funds carried out in 2004-2006 programming period served as an important learning experience. While only an ex ante evaluation was compulsory in view of the relatively short programming period, a number of Hungarian institutions took the initiative to carry out mid-term⁵⁶ and thematic evaluations. In the period 2004-2006 some 8 evaluations related to Structural Funds were carried out (excluding ex ante evaluations of 2007-2013). The list of evaluations commissioned is included in annex 1 to this report. There were variations between institutions as regards the number of evaluations commissioned. The Ministry of Economy (MA of the Economic Competitiveness OP) and Ministry of Social Affairs and Labour (MA of Human Resources Development OP) were among the most active MAs. The MA of the Economic Competitiveness OP commissioned a mid-term evaluation of while the MA of the Human Resources Development OP commissioned interim evaluation consisting of three evaluation assignments (concerning monitoring indicators, horizontal issues at a project level and the contracting process). The NDA⁵⁷ commissioned an ex ante evaluation of the National Development Plan for 2004-2006 and an evaluation of the CSF institutional framework and resource allocation mechanisms. For the period 2007-2013, NDA decided to conduct an ex ante evaluation of the NSRF, which was not compulsory but recommended for large Convergence Objective Member States.⁵⁸ NDA was also responsible for evaluation of the OPs.⁵⁹ Ex ante evaluations were conducted for the group of OPs, except regional OPs, which were evaluated on an individual basis. The relevant MAs were involved in the ex ante process via steering groups.

Information on financial resources for evaluation in 2004-2006 programming was not available in respect of each OP. The budget for ex ante evaluations was approximately 710 000 euros. In 2007-2013 approximately 1.3 million euros are foreseen for OP level evaluations. It is worth mentioning that the general view of survey and interview respondents (both civil servants and external evaluators) is that financial resources allocated to evaluation function are adequate.

⁵⁶ The mid-term evaluation being not mandatory in the newly acceded countries.

⁵⁷ The *ex ante* evaluation of the National Development Plan was commissioned by the Office of the National Development Plan and EU Support of Hungary.

⁵⁸ European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013. Indicative Guidelines on Evaluation Methods: Ex Ante Evaluation. Working Document No.1. August 2006. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd1_exante_en.pdf

⁵⁹ For some OPs *ex ante* evaluation was initiated by their respective MAs and not NDA: i.e. in case of Economic Development OP and Environment and Energy OP.

In terms of evaluation staffing resources, the Department of Evaluation and Methodology under NDA is made of thirteen persons. As regards the MAs of the OPs, there are one to two persons responsible for evaluation. Most of them have also other tasks apart from evaluation (such as monitoring) so these levels are not seen as adequate. The current view of the NDA is that staff numbers will have to increase to two-four persons, depending on the size of the OP.

Referring to the issue of expertise, most of the staff at the Department of Evaluation and Methodology have a background in economics. The staff have no prior experience in carrying out evaluations, reflecting the reality that evaluation practice was limited before accession. However, considerable experience in commissioning evaluations, drafting terms of references (as discussed below, all terms of references undergo the Department's quality check), steering projects and exercising quality control has been accumulated over 2004-2006 period.

The issue of evaluation capacity within the administration is of particular importance in view of the planned staff expansion at the MAs, the evaluation planning requirements and the ambition to make evaluation a genuine management tool. At present there is no document setting out evaluation capacity building measures or directions. The options for training in evaluation are limited in Hungary. There were some evaluation capacity building events financed by pre-accession programmes and carried out as part of preparation for EU membership. Annual evaluation conferences take place at the end of the year. To date two such events were organised: in 2005 and in 2006. Both civil servants and evaluators are invited to participate in these conferences. There is also a forum established for development policy makers under the title of „Academy for Development Policy“. While it is not confined to evaluation and deals with broader issues of development policy, thematic conferences and training sessions are organised, including those related to evaluation (e.g. an event dedicated to ex ante evaluation was organised).

An interesting practice to mention is cooperation between civil servants of the Visegrad group, of countries (Poland, Hungary, Slovak and Czech Republics) in the field of Structural Fund evaluation. While at present there is a pause in cooperation, one of the interviewees mentioned it as a useful vehicle for learning and sharing experience which they would like to see continued.

Efficiency of Evaluation Management

As mentioned above, the authorities have already accumulated certain evaluation management experience, even though it is not equally spread across institutions. Evaluation guidelines prepared by NDA are intended to assist institutions in the evaluation management process.⁶⁰ Terms of reference for individual evaluation projects are drafted by the Department of Evaluation and Methodology. Then they are passed on to the relevant MAs for sectoral input. Current experience shows that drafting and agreeing on terms of reference takes approximately two to four months (in some cases – up to six months). In future the intention is to develop some standardised form or core of terms of reference, which would facilitate the drafting process. While the detailed analysis of all the terms of references for evaluation assignments was not feasible for the purposes of this report, it is worth noting that external evaluators regard them as of good quality.

After terms of reference are agreed, the length of the public procurement process depends on the size of evaluation project and the tendering procedure. Timing issues are set in the Hungarian Act on Public Procurement. Some Structural Fund evaluations were contracted by an open procurement procedure, which takes longer, and some by a restricted procedure. In an open procedure all interested parties may submit offers. In a restricted procedure only companies invited by the Contracting Authority may participate (the invitation to tender must include minimum five tenderers). Service contracts may be awarded on the basis of one out of two criteria: either the lowest price or the most economically advantageous tender in terms of price and quality. The common practice is to rely on the latter criterion. Typically the price-quality ratio is 40:60 or 50:50.

⁶⁰ Existent guidelines are largely based on EU guidelines.

Ad hoc steering committees are established to oversee individual evaluation projects (e.g. ex ante evaluations). The MAs also set up evaluation steering groups, which consisted of representatives from the MA, intermediate bodies and relevant ministries and which met to discuss such issues as drafting terms of reference, assessing the quality of reports, etc. There are no pre-set quality criteria against which the evaluation reports are assessed. In case of ex ante evaluation there was a checklist developed with quality criteria which the report had to conform in order to be accepted. The checklist consisted of formal and content related criteria.

An important aspect of the management of evaluation activity is an evaluation plan. As stipulated by the *Council Regulation No 1083/2006*, Member States may draw up under the Convergence objective an evaluation plan presenting indicative evaluation activities which the Member State intends to carry out in different phases of the implementation.⁶¹ The main purpose of establishing an evaluation plan is to provide an overall framework for the on-going evaluation and ensure that it is effectively used and integrated as a management tool during the implementation phase.⁶² In 2007-2013 Structural Fund evaluation in Hungary will take place on the basis of such plans which are in the course of preparation at both national and programme level.

The evaluation plan at the NSRF level, initially covering the three-year period 2007-2009, is being drafted by the Department of Evaluation and Methodology under NDA. It will have to be agreed on by institutional stakeholders. The intention is to have the plan approved at the level of the NDA President. The draft evaluation plan sets out evaluation objectives, types, stakeholders and specific evaluation assignments included into the plan according to pre-defined criteria (such as size, budget, complexity, innovativeness of the intervention, implementation progress and etc). The following issues will be the subject of evaluation in an on-going evaluation system in 2007-2013: system of indicators and evaluation reports, individual interventions, action plans⁶³ (ex ante and final evaluation), OPs and the NSRF (mid-term and ex post evaluation).⁶⁴ The evaluation plan will be updated on annual basis to ensure the necessary flexibility.

The evaluation plans at OP level will cover a period of 3 years and will be revised on an annual basis. The OP evaluation plan will be submitted by the MA of a the respective OP, with the agreement of the Department of Evaluation and Methodology, to the OP Monitoring Committee, which will consider and approve the annually updated evaluation plan. As standard text in the OPs suggests, the following evaluations will be included in the OP evaluation plan:

- mid-term revision of the OP strategy and implementation system (2009-2010);
- ex post evaluation of the OP (2015-2016);
- annual operational evaluation of action plans (2008, 2010, 2012, 2014);
- ex post evaluation of action plans (2009, 2011, 2013, 2015);
- ex ante evaluation of the contribution of action plans, individual interventions and major projects to the implementation of objectives of the OP, and to the enforcement of horizontal policies (2006, 2008, 2010);
- comprehensive evaluation of the OP with regard to horizontal themes (2008, 2010, 2012).

Planning of the evaluation activity is a new experience in Hungary. It is expected to contribute to better management of demand for evaluation during the entire programming period.

Quality of Monitoring Systems

⁶¹ Council Regulation No. 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999

⁶² European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf

⁶³ Documents similar to Programme Complements in the previous programming period.

⁶⁴ New Hungary Development Plan for 2007-2013. - P.159.

Both administrations and external evaluators acknowledge that there is still room for improvement as regards precision, quality and timeliness of monitoring data, especially physical data. This is also confirmed by the related evaluation reports (e.g. interim evaluation of HRDOP: study on indicators).⁶⁵ Concerns were expressed regarding the definition of indicators, their aggregation from project to higher levels and timeliness of loading data into the system, etc. It was not uncommon for evaluators to collect the necessary monitoring data during the evaluation process via such data collection methods as surveys or interviews.

(4) Supply Side

Having discussed the demand side, the analysis below focuses on the supply-side analysed under the two criteria, namely:

- Quality of socio-economic data resources; and
- Availability and quality of evaluation expertise.

Quality of Socio-economic Data Sources

Availability of reliable external data on key socio-economic variables is essential for evaluation of Structural Fund programmes with a wide range of policy objectives. In general, the quality and timeliness of contextual, socio-economic indicator data (e.g., GDP, employment, unemployment, R&D investment) at national and NUTS2 level is felt to be adequate in Hungary. The latest year for which most data is available is 2006. At the same time external evaluators noted that sector specific data is more difficult to obtain.

Availability and Quality of Evaluation Expertise

Structural Fund authorities in Hungary primarily rely on external services for the evaluation of Structural Fund programmes. Several considerations account for this. Firstly, internal evaluations would require more resources in terms of human resources and also adequate expertise. Secondly, the intention of the authorities is to encourage further development of the market and the quality of services.

Demand for evaluations in the 2004-2006 programming period stimulated the emergence of the evaluation market and competition between companies. Some 10 companies are active in providing evaluation services, while the total number of companies in the evaluation market is over 20. A few companies are based in the regions. Foreign companies are involved in providing evaluation services (e.g. the first ex ante evaluation of NDP 2004-2006 was carried out by a foreign company) but to a limited extent, and in most cases as partners. With some exceptions universities and research institutions are not active in the Structural Fund evaluation area but a number of academic institutions are engaged in broader research activity related to cohesion policy (e.g. Centre of Regional Studies of the Hungarian Academy of Sciences, which has also experience in ex ante evaluation of Regional Development OP).

The quality of evaluation reports drafted by external service providers was regarded by most survey respondents as being fair or good. The view expressed by several interviewees was that there was still a lack of strong theoretical foundation or theory-driven evaluation. Up to now the tendency has been that of a more “practical” approach. Greater involvement of researchers from academic or research institutions could be one way of addressing this issue and bringing a stronger theoretical and methodological orientation. Here it is worth mentioning that the authorities (i.e. NDA) took an initiative to develop a macroeconomic model which allowed for the estimation of the Structural Fund effects at national, sectoral and regional levels.

⁶⁵ Study „Evaluation of Indicators“. Part of the mid-term evaluation of the Human Resource Development Operative Programme. Executive summary in English. – Ex Ante Ltd. – 2005.

As regards the availability of education and training in evaluation of Structural Fund programmes, the choice is felt to be limited in Hungary. As already mentioned, evaluation conferences that bring together representatives from public administration and external evaluators have been organised on annual basis for the last two years. During these conferences new evaluation reports, evaluation techniques and good practices have been discussed. In general, the view is that strengthening of external evaluation capacity is a matter for the market rather than for the public administration. However, the market has not yet responded sufficiently to this need. But, as it will be discussed in the following sections, steps have been taken to establish the Hungarian evaluation society, which is intended to address, *inter alia*, training issues.

(5) Dissemination and Utilisation

With reference to public access to evaluation results, summaries of many evaluation reports are available on a webpage <http://www.fejlesztéspolitika.gov.hu> where one section is dedicated to evaluation of Structural Funds. External evaluators note that increased accessibility to full evaluation reports could further contribute to better awareness of evaluations commissioned, methodologies used and facilitate professional discourse.

As regards the utilisation of evaluation results, in 2004-2006 programming period there were no formal procedures laid down as to how evaluation reports and recommendations should be dealt with. These were discussed and taken into consideration by steering groups. The Monitoring Committee was informed about the evaluations carried out and could take decisions concerning the utilisation or dissemination of evaluation results. While case study evidence does not allow for clear-cut conclusions with regard to utilisation of evaluation results, a mixed picture seems to emerge. Evaluators feel that evaluations have had certain impact and there is some evidence of the implementation of evaluation recommendations (e.g., monitoring indicators of the Human Resource Development OP were revised as a result of evaluation on indicator system). However, as noted by several interviewees, the timing of some evaluations was such as to limit their utilisation (as their relevance diminished due to the end of the programming period and changes in the overall Structural Fund administration system). The impact of evaluation is, for obvious reasons, most noticeable in case of *ex ante* evaluations. It was felt that *ex ante* exercise had a positive influence on the programming process, not least due to the interactive nature of approach taken. In general, the survey results and interviews suggest that influence of *ex ante* evaluation was significant on indicators and situation analysis, with limited impact on the distribution of resources – a conclusion which is similar to findings in other Member States.

It is noteworthy that for the 2007-2013 period formal arrangements for addressing evaluation findings and tracking agreed recommendations are planned. The role of steering groups will be extended - i.e. steering groups will be delegated a responsibility for following-up the implementation for evaluation recommendations after the approval of an evaluation report.

(6) Degree of Institutionalisation of Evaluation in the Wider Governance System

This section considers the various criteria specified under the institutionalisation dimension of the Evaluation Capacity Development Benchmarking Framework as follows:

- The role of evaluation in the national budgetary process;
- Role of evaluation in the formulation of sectoral strategy and policy making;
- The existence of mandatory evaluation requirements;
- The public policy research base;
- National evaluation society;
- Influence of EU evaluation requirements; and
- The role of civil society.

In accordance with the Hungarian Act on Legislation, institutions involved in rule-making and implementation are obliged to assess the effects of legal regulations and make known any obstacles

to their application, ineffectiveness or implementation difficulties. While evaluations are used to some extent in the national budgetary and policy-making process, evaluation is still mostly related to and developing within the framework of the Structural Funds. As highlighted in 2007-2013 State Reform OP, policy-making process still needs to be improved by paying more attention to impact assessment and evaluation of draft legislation as well as subsequent implementation.⁶⁶

It must be recognised that a certain evaluation tradition existed in Hungary before its accession to the EU, mainly in the field of environmental protection and regional development. For example, VÁTI, a public non-profit company for regional development and town planning, has been dealing with evaluation issues for some fifteen to twenty years. There are a number of public policy research institutions involved in public policy analysis. For example, Centre for Policy Studies (CPS), based at the Central European University, carries out public policy research projects, publishes research papers, develops material for teaching and training, offers master's degree in public policy and organises conferences. This is aimed at improving policy-making capacity in Hungary. Social Research Centre (TARKI) is a well-known private research organisation specialising in social policy studies and analysis of the social consequences of economic policies.⁶⁷ The Foundation for Market Economy focuses on economic policy. It carries out research and attempts to use results for the policy-making and training purposes. The aforementioned VÁTI and the University of Pecs are active in research on regional development policy.

Focusing on the question of the professional evaluation community, it is important to note that steps have already been taken to establish Hungarian evaluation society. This year initial meeting was held and working group for drafting the statute was set up.

Research and evaluation findings are not widely used by civil society actors. The role of civil society and other players in the evaluation sphere has been very limited in Hungary to date. As noted by one of the interviewees, the interest in evaluation findings *inter alia* depends on a type of evaluation. For example, it is expected that results of *ex post* evaluation of Structural Fund programmes would attract more attention than *ex ante* evaluations, which are rather specific and strongly related to the drafting process.

(7) Conclusions

At the time of writing preparations for the 2007-2013 programming period are on-going in Hungary. While detailed arrangements for evaluation are still to be put into practice, some general conclusions concerning evaluation of Structural Fund programmes may already be presented taking into account the initial experience gained in the period of 2004-2006. Considerable attention is paid to Structural Funds evaluation in Hungary. This is evidenced by the number of evaluations commissioned, evaluation guidelines drafted, financial resources devoted, the plans to increase the number of evaluation staff and initiatives taken to develop specific methodologies (i.e. macro-economic model), and design procedures for tracking agreed evaluation recommendations. To date the Department of Evaluation and Methodology under NDA has been performing an important role as evaluation coordinator, promoter and as a centre for methodological excellence. At the same time evaluation capacities at the OP level have been put in place and are planned to expand.

One of the potential challenges facing authorities in the 2007-2013 programming period is ensuring consistency and efficient coordination of evaluation needs between different levels and programmes and achieving broad acceptance of evaluation as a useful management tool (taking into account that so far interest in evaluation has been uneven across different MAs.) On the one hand, the coordination role of the Department of Evaluation and Methodology will assist in addressing this challenge. At the same time it is important that MAs take a leading role in defining the evaluation needs of their programmes. The OP evaluation plans and the new formalised mechanisms for

⁶⁶ State Reform OP.

⁶⁷ About TARKI // <http://www.tarki.hu/en/about/>

addressing evaluation results may assist in that regard. Another issue that authorities will need to handle is that of maintaining and further developing evaluation capacity, especially at OP level. If the number of evaluation staff at the MAs increases as foreseen, it is important to address their capacity building and training needs.

The successful implementation of the on-going evaluation system will also require indicator and monitoring systems that are capable of providing reliable information to managing authorities in addition to providing the essential data inputs to evaluation projects.

As regards the supply side, a competitive evaluation market is gradually developing. To date private companies have dominated the provision of evaluation services with academic or research institutions involved only to a limited extent. With little prior tradition of evaluation in Hungary, external evaluation capacity has been developing in what might be termed as “learning by doing” manner. Steps taken to establish an evaluation society are promising as this will potentially contribute to the emergence of professional discourse and learning. Centralised access to and wider dissemination of evaluation reports could further contribute to the spread of good practice.

(8) Key Learning Points

With regard to evaluation capacity benchmarking framework

- With reference to the application of the evaluation capacity benchmarking framework, it provided a useful basis on which to organise the field work and to structure this case study report. Some difficulties were encountered in gathering information against some indicators. Hard quantitative data was not available for the indicator number 6 relating to evaluation budget in 2004-2006 programming period. Detailed assessment of quality of all terms of reference was not feasible, thus, we could rely only on the general views expressed by interviewees and survey respondents. The same caveat applies to the assessment of the quality of evaluation services.

From Hungarian experience

- As regards lessons from the Hungarian experience, the pro-active stance of evaluation unit at national level acted as a significant driver for demand and the development of evaluation capacity in the Structural Fund administration in Hungary. The short 2004-2006 programming period was used to gain evaluation experience with institutions commissioning not only compulsory, but also non-mandatory Structural Fund evaluations. The demand for evaluation stimulated the emergence and development of evaluation market which confirms the view that demand for evaluation is an essential starting-point in building evaluation capacity. The establishment of an evaluation society can be expected to further contribute to external evaluation capacity development.
- While the role of evaluation in national policy formulation process is limited and develops within the framework of Structural Funds, structural funding (State Reform OP) will be used in 2007-2013 programming period to support improvements of the national policy-making process, not least by strengthening the attention to evaluation and impact assessment.

Bibliography

1. About National Development Agency. // <http://cordis.europa.eu/erawatch/index.cfm?fuseaction=org.document&uuid=7D87D84D-DB9A-B9CF-49E2A7A02BFBAF4D>
2. Council Regulation No. 1083/2006 of 11 July 2006 laying down general provisions on the European Regional Development Fund, the European Social Fund and the Cohesion Fund and repealing Regulation (EC) No 1260/1999
3. European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013. Indicative Guidelines on Evaluation Methods: Ex ante Evaluation. Working Document No.1. August 2006.. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd1_exante_en.pdf
4. European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. Indicative guidelines on evaluation methods: evaluation during the programming period. Working paper No.5. April 2007. http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd5_ongoing_en.pdf
5. Evaluation system descriptions. Visegrad Evaluation Meeting II. – Zakopane, September 28-30, 2005.
6. Hungarian Act on Public Procurement No. CXXIX, 2003.
7. Hungarian National Development Plan 2004-2006. – 28 March 2003.
8. Hungary. Cohesion Policy 2007-13. // http://ec.europa.eu/regional_policy/atlas2007/fiche/hu_en.pdf
9. Operational programmes in Hungary in 2007-2013. // http://www.nfu.gov.hu/index.nfh?r=&v=&l=&d=&mf=&p=umft_opprog
10. The New Hungary Development Plan. National Strategic Reference Framework of Hungary 2007-2013. Employment and Growth. *Date of the decision of the European Commission: 7 May 2007*

Annex to the Hungary case report: List of Structural Fund evaluations commissioned in Hungary during 2003-2007

Year	Title of Structural /Cohesion Fund related evaluation assignment	Contracting Institution	Company /Organisation
2003	<i>Ex ante</i> evaluation of the National Development Plan (2004-2006)	Office of the National Development Plan and EU Support (currently called: National Development Agency)	ÖIR-Managementdienste GmbH
2005	Mid-term evaluation of the Economic Competitiveness Operational Programme (2004-2006)	Management Authority of the Economic Competitiveness Operational Programme	Consortium of KPMG-Hungary and PWC-Hungary
2005	Guidelines and trainings on the ex post evaluation of Cohesion Fund projects	Office of the National Development Plan and EU Support (currently called: National Development Agency)	PricewaterhouseCoopers Ltd.
2005	Guidelines, model and trainings on Cost-Benefit-Analyses	Office of the National Development Plan and EU Support (currently called: National Development Agency)	PricewaterhouseCoopers Ltd.
2005	Mid term evaluation of the HRDOP project development and project support	Ministry of Employment and Labour	PricewaterhouseCoopers Ltd.
2005	Mid term evaluation of the National Development Plan (2004-2006)		National Development Office – Unit for Analysis, Evaluation and Modelling
2005	Evaluating the institutions implementation system of the Community Support Framework (2004-2006)	National Development Office (currently called: National Development Agency)	Consortium of Ex Ante Ltd and HBF Hungaricum Ltd
2005	Evaluating the resource allocation mechanisms of the National Development Plan (2004-2006). In-depth analysis of 7 micro-regions (NUTS IV)	National Development Office (currently called: National Development Agency)	HBF Hungaricum Ltd
2005	Comprehensive evaluation of the Hungarian Employment Policy in terms of the European Employment Strategy	Ministry for Employment and Social Affairs	Institute of Economics – Hungarian Academy of Sciences
2006-2007	Impact assessment of the Hungarian Operational Programmes in terms of 5 horizontal issue (expected impacts on: budget deficit, long-term growth, long-term employment, situation of disadvantaged groups, allocation of welfare among social groups)	National Development Agency	Consortium of Kopint-Datorg and TÁRKI
2006-2007	Strategic Environmental Assessment of the National Reference Framework and Operational Programmes for the 2007-2013 period	National Development Agency	Respect Ltd
2006-2007	<i>Ex ante</i> evaluation of the National Strategic Reference Framework (2007-2013)	National Development Agency	Consortium of HBF Hungaricum Ltd and Ex Ante Ltd
2006-2007	<i>Ex ante</i> evaluation of Social Renewal Operational Programme (2007-2013)	National Development Agency	Consortium of Compudoc and ForEnviron Ltds

2006-2007	<i>Ex ante</i> evaluation of Social Infrastructure Operational Programme (2007-2013)	National Development Agency	Consortium of Compudoc and ForEnviron Ltds
2006-2007	<i>Ex ante</i> evaluation of Electronic Administration Operational Programme (2007-2013)	National Development Agency	Consortium of Compudoc and Ex Ante Ltds
2006-2007	<i>Ex ante</i> evaluation of State Reform Operational Programme (2007-2013)	National Development Agency	Consortium of Compudoc and Ex Ante Ltds
2006-2007	<i>Ex ante</i> evaluation of Economic Development Operational Programme (2007-2013)	National Development Agency	Consortium of KPMG Hungary and PWC Hungary Ltds
2006-2007	<i>Ex ante</i> evaluation of Environment and Energy Operational Programme (2007-2013)	National Development Agency	Öko Rt
2006-2007	<i>Ex ante</i> evaluation of Transport Operational Programme (2007-2013)	National Development Agency	Consortium of COWI, Ex Ante and Okean Ltds
2006-2007	<i>Ex ante</i> evaluation of South-Great Plain Operational Programme (2007-2013)	National Development Agency	Consortium of Terra Studio and KSZI Ltds
2006-2007	<i>Ex ante</i> evaluation of South-Transdanubia Operational Programme (2007-2013)	National Development Agency	Consolidus Co. Ltd
2006-2007	<i>Ex ante</i> evaluation of Northern Hungary Operational Programme (2007-2013)	National Development Agency	Pillars Ltd
2006-2007	<i>Ex ante</i> evaluation of Western Transdanubia Operational Programme (2007-2013)	National Development Agency	Pillars Ltd
2006-2007	<i>Ex ante</i> evaluation of The North Great Plain Operational Programme (2007-2013)	National Development Agency	Megakom Ltd
2006-2007	<i>Ex ante</i> evaluation of Central Transdanubia Operational Programme (2007-2013)	National Development Agency	Consolidus Co. Ltd
2006-2007	<i>Ex ante</i> evaluation of Central Hungary Operational Programme (2007-2013)	National Development Agency	Consortium of Terra Studio and KSZI Ltds
2006-2007	<i>Ex ante</i> evaluation of Implementation Operational Programme (2007-2013)	National Development Agency	Consortium of Compudoc and Ex Ante Ltds
2007	<i>Ex ante</i> evaluation of the action plans of the Economic Development Operational Programme (EDOP) 2007-2013	National Development Agency, Managing Authority of the EDOP	CEDEC Ltd.

Report on Lithuania Case Study

(1) Introduction and Methodology

This report presents the results of the case study of Structural Fund evaluation capacity in Lithuania. The case study is based on a review of relevant documentation and semi-structured face-to-face interviews carried out in Vilnius in June and July 2007. The full list of documents reviewed is set out in the bibliography. 9 persons were interviewed for the purposes of the case study as follows:

- A joint interview was carried out with 2 representatives of Evaluation Unit at the Ministry of Finance, which performs the functions of SPD Managing Authority in 2004-2006 and National Coordinator in 2007-2013.
- One interview was conducted with three representatives of the Ministry of Social Security and Labour, an Intermediate Body under the SPD 2004-2006⁶⁸ and the Human Resource Development Operational Programme 2007-2013.
- A joint interview was carried out with 2 representatives of the Ministry of Transport and Communications, Intermediate Body under the Economic Growth Operational Programme of 2007-2013.
- Two interviews were conducted with representatives of consulting companies with experience in Structural Fund evaluations.

The results of the evaluation capacity benchmarking survey were also taken into consideration.⁶⁹ The documentary analysis, interviews and survey provided useful information on the various components included in the Evaluation Capacity Development Benchmarking Framework, structured around four core dimensions as follows:

1. Demand, which refers to the commissioning side of the evaluation process;
2. Supply;
3. Dissemination and utilisation of evaluation outputs;
4. Institutionalisation dimension which aims at capturing a wider administrative culture (i.e., outside the Structural and Cohesion Fund administration).

In the report below some general information on the Structural Funds in Lithuania is firstly presented. This is then followed by the analysis of evaluation capacity according to the aforementioned dimensions of the evaluation capacity development benchmarking framework.

(2) Background Information: Structural Funds in Lithuania

Lithuania became a member of the European Union on 1 May 2004. During the period 2004-2006 Lithuania was eligible for Structural Fund support under Objective 1. A Single Programming Document (SPD) for this period was prepared and funded by approximately 895 million euros of Structural Funds.⁷⁰ The Ministry of Finance was nominated as a Managing Authority (MA). The of intermediate body functions were performed by seven line ministries and one committee under the Government. A further six implementing bodies were responsible for project administration and monitoring under certain SPD measures.

⁶⁸ Also responsible for evaluation of Equal.

⁶⁹ In total 8 representatives from administration and 4 market participants completed the on-line survey questionnaire.

⁷⁰ This amount has to be supplemented with approximately 310 million euros of national investment.

For the 2007-2013 period Lithuania is classified as a single Convergence objective region. Lithuania has been allocated approximately 6.7 billion euros of Structural and Cohesion Funds,⁷¹ which will be invested in accordance with National Strategic Reference Framework (NSRF)⁷² and operational programmes (OPs). Lithuania has translated the objectives and priorities of the NSRF into the following four OPs:⁷³

No.	Operational Programme	EU Fund
1.	Development of Human Resources	ESF
2.	Economic Growth	ERDF and Cohesion Fund
3.	Cohesion Promotion	ERDF and Cohesion Fund
4.	Technical Assistance	ESF

The Ministry of Finance has been designated as a Managing Authority of all OPs. It will also carry out functions of a National Coordinator. The key responsibilities of a National Coordinator include the overall strategic coordination of EU cohesion policy and structural assistance in Lithuania and monitoring the performance of management and control systems.

Having presented the background information on Structural Funds in Lithuania, the following sections focus on evaluation capacity analysed on the basis of the criteria and indicators proposed in the evaluation capacity development benchmarking framework.

(3) Demand Side

The analysis in this section is presented under 4 criteria included under the demand-side dimension of the evaluation capacity development benchmarking framework, i.e.:

- The integration, location and organisation of the Structural Funds evaluation function;
- Evaluation resources;
- Efficiency of evaluation management, and
- Quality of monitoring systems.

Integration, Location and Organisation of Evaluation Function

Lithuania adopted a centralised approach to the organisation of the Structural Funds evaluation function in the 2004-2006 period and will retain this in 2007-2013. The Ministry of Finance in its capacity as the MA was responsible for management of of SPD evaluation in 2004-2006.⁷⁴ Intermediate bodies bore responsibility for evaluations at SPD measure level and were required to participate in the work of Evaluation Management Group (EMG).⁷⁵ The EMG was set up in 2005 and served as a vehicle for co-ordinating evaluation of the SPD including analysing terms of reference for evaluation projects, discussing draft evaluation reports and the making of recommendations to authorities responsible for evaluation at a project level.⁷⁶ The EMG consisted of representatives of the MA, intermediate bodies and independent evaluation experts, with the latter invited to the meeting depending on the issues under discussion.

As regards the regulatory basis, the Government's Resolution "*On the Administration of Assistance from the EU Structural Funds in Lithuania*" of 2003 was a key document which delineated the

⁷¹ That will be supplemented with an estimated 1.5 billion euros of national investment bringing the total sum for cohesion policy 2007-13 in Lithuania to about 8.2 billion euros. // Lithuania. Cohesion Policy 2007-2013. http://ec.europa.eu/regional_policy/atlas2007/fiche/lt_en.pdf

⁷² Approved by the European Commission on 26 April 2007.

⁷³ Lithuania. Cohesion Policy 2007-2013. // http://ec.europa.eu/regional_policy/atlas2007/fiche/lt_en.pdf

⁷⁴ Also evaluation of Equal and Interreg Community Initiatives. // Resolution of the Government of the Republic of Lithuania No. 1166 of 11 September 2003 "*On the Administration of Assistance from the EU Structural Funds in Lithuania*"

⁷⁵ Guidelines for Evaluating EU programmes in Lithuania. – 2005.

⁷⁶ Guidelines for Evaluating EU programmes in Lithuania. – 2005.

division of evaluation responsibilities between institutions. The procedures manual for the Ministry of Finance (in its capacity as MA) set out the process for organising and coordinating evaluation work. However, the organisation of the evaluation function at intermediate bodies was insufficient. As noted by the authorities, in many instances department statutes and job descriptions did not refer to Structural Fund evaluation as a function, which meant that evaluation represented an additional activity for personnel working with structural assistance and was relegated to a secondary position.⁷⁷ While regulations *per se* might not be expected to transform evaluation practice radically, they represent an important step towards the institutionalisation of evaluation.

At the time of writing, a lot of the arrangements for the evaluation function in 2007-2013 were in the course of development in Lithuania. The body responsible for the overall organisation and coordination of OP evaluation is the National Coordinator (Ministry of Finance). With regard to Structural Fund evaluation, the National Coordinator bears responsibility for the overall coordination and organisation of evaluation function, drafting of the evaluation plan, coordinating its implementation and setting up an Evaluation Coordination Group. These functions will be carried out by the Evaluation Unit of the Ministry of Finance, established in April 2006.

Intermediate bodies are obliged to participate in carrying out Structural Fund evaluation functions. The programme level evaluations of *operational* nature (to be initiated when the monitoring reveals a significant departure from the goals initially set or where proposals are made for the revision of operational programmes) will be organised by the MA, in close cooperation with intermediate bodies. *Strategic* evaluations will be carried out on the initiative of the MA or intermediate bodies. Intermediate bodies shall also carry out *ex ante*, mid-term or *ex post* evaluations of major projects.

For coordination purposes the Evaluation Coordination Group will be appointed. At the time of writing it has not been decided whether this group would replace the EMG or function alongside it (with the EMG dealing with the evaluation of SPD and Evaluation Coordination Group dealing with evaluation of OPs). Similar to the EMG, the Evaluation Coordination Group will consist of the representatives of National Coordinator, MA, intermediate bodies and guests (e.g. external evaluation experts, socio-economic partners, European Commission) to be invited to the meetings depending on the issue under discussion.

The Monitoring Committee will approve the evaluation plan and consider evaluation reports.

Evaluation Resources

Having presented institutional arrangements for evaluation of Structural Fund programmes in Lithuania, this section considers the amount of evaluation activity and resources (human and financial) allocated to it.

Apart from the evaluation of pre-accession support (Phare, ISPA, SAPARD), which started in 1999 and was carried out according to specific requirements, the first experience in evaluation of the Structural Fund programme was the *ex ante* evaluation of the Lithuanian SPD for 2004-2006 and its Programme Complement document.⁷⁸ There was no requirement to carry out a mid-term evaluation of the SPD. Despite the absence of mandatory requirements, the national authorities launched several *ad hoc* evaluations. (A list of these evaluations is included in Annex 1 to this case study report.) The need for specific evaluations became apparent in the course of SPD implementation, and in planning investments for the 2007-2013 programming period.

On the whole, the attention paid to evaluation varied across Structural Fund institutions. The MA was the most active in commissioning evaluations: to date five evaluations have been initiated (excluding mandatory *ex ante* evaluations). These concerned such national-level issues as the SPD implementation system, Structural Fund investment areas for 2007-2013, Structural Fund publicity and information system, implementation of horizontal priorities and Structural Fund employment

⁷⁷ Evaluation capacity analysis carried out by the Ministry of Finance in April-May 2007.

⁷⁸ Also Interreg and Equal.

effects.⁷⁹ The Ministry of Finance was authorised to organise a (single) ex ante evaluation for all OPs for 2007-2013. The ex ante evaluation project was steered by the EMG and recommendations were discussed by the Commission responsible for programming the EU Structural Funds in Lithuania. Finally, it is worth noting that one small-scale internal evaluation was conducted by the Evaluation Unit in April-May 2007. This took the form of a survey on evaluation capacity in the Structural Fund administration in Lithuania and attempted to assess institutional and administrative dimensions of evaluation capacity and use the results for planning evaluation capacity building activities in 2007-2013.

As to other institutions, five intermediate bodies took the initiative to commission evaluations at a measure (e.g., of the SPD measure 1.1 "*Improvement of Accessibility and Service Quality of Transport Infrastructure*") or sectoral level (e.g., analysis of human resource sector and recommendations for investment strategy in 2007-2013).⁸⁰ It can be mentioned that these were institutions where employees have been delegated the responsibility for Structural Fund evaluation and have been working in this area for a longer period. Even though implementing bodies are not delegated responsibility for Structural Fund programme evaluation in Lithuania, several implementing bodies took initiative and commissioned evaluations concerning issues, relevant to them – such as, for instance, beneficiary attitudes towards the process of preparation of applications for structural funding.

A survey carried out by the Evaluation Unit revealed that where institutions had not commissioned any evaluation, the main reasons cited were the absence of mandatory requirements or a pressing need. Some institutions noted that the volume of evaluations commissioned by the Ministry of Finance was sufficient and that additional evaluations were required. Finally, a few intermediate bodies did not conduct any evaluations as it was felt too early to evaluate the results of the SPD implementation. A lack of financial resources was never quoted as a reason for not conducting evaluations.⁸¹ However, several interviewees indicated that the procedure for obtaining financing for evaluation projects was an obstacle (as in order to receive financing for evaluation projects, institutions had to submit applications for technical assistance).

Evaluations of the SPD for 2004-2006 were financed from the SPD technical assistance measure 5.2 "*Programme Publicity and Evaluation*".⁸² Application for technical assistance could include both publicity and evaluation activities, which complicates estimation of the amount of financial resources dedicated specifically to evaluation activities, especially considering that not only evaluations in a strict sense but also feasibility studies and other kinds of research were financed. Calculations made by the Evaluation Unit suggest that approximately 0.9 million euros were spent on evaluation in 2004-2006, which equals to 0.08 percent of total Structural Funds assistance. In the period of 2007-2013 approximately 9 million euros are foreseen for evaluation, which equates to 0.06 percent of total structural assistance for Lithuania.

While available financial resources for evaluations were adequate, survey, interviews and evaluation capacity analysis carried out by the Ministry of Finance indicate that current human resource levels are insufficient, in particular at the intermediate bodies. In most intermediate bodies one person, or, in a few cases, two persons, work on evaluation of structural assistance. As already mentioned, in many cases evaluation is not a full-time but a supplementary responsibility and is often not included in job descriptions, which hinder its effective implementation and the development of "ownership". In the meantime, a few institutions are already planning to employ full-time personnel to deal with evaluation issues in 2007-2013 programming period.⁸³

⁷⁹ At the time of writing, some of the evaluations mentioned were still in progress.

⁸⁰ List of evaluations commissioned is provided in Annex 1 to this case study report.

⁸¹ Evaluation capacity analysis carried out by the Ministry of Finance in April-May 2007.

⁸² Except for ex ante evaluation of the SPD, which was financed by PHARE.

⁸³ For example, Ministry of Transport and Communications.

Within the MA, initially one full-time member of staff was responsible for managing evaluation. With the establishment of the Evaluation Unit in April 2006, four full-time positions were created. Currently three persons work in the Evaluation Unit. In general, this is felt to be adequate for carrying out coordination functions, but not internal evaluations. As regards education of the Evaluation Unit personnel, they have a background in social sciences (business administration and the EU studies, sociology and law). None of the staff members had prior experience in carrying out evaluations, which reflects the fact that evaluation capacity started to develop mostly with the introduction of the Structural Funds in Lithuania.

In order to increase awareness and knowledge of the Structural Fund evaluation both within the administration and the wider market, several capacity building projects were implemented in Lithuania. Special training on Structural Fund programme evaluation was organised in 2004 under the Phare SPP I and SPP II. Approximately 30 participants, including civil servants and external experts, took part in the 21-hour training “Evaluation of the European Union Structural Funds”.

Another noteworthy initiative was the evaluation capacity building project initiated by the MA and carried out in 2005-2006. The project consisted of 3 types of activities: (i) Evaluation training activities; (ii) Preparation of methodological guidance documents and (iii) Awareness raising measures. These measures were aimed at building evaluation capacity both in the Structural Fund administration and evaluation market. The Box 1 below describes project activities in more detail.

Box 1. Project on Strengthening Evaluation Capacity in Lithuania, 2005-2006

For the purpose of evaluation capacity building project an assessment of training needs in evaluation was first conducted. In 2005 more than a half of surveyed civil servants working in the Structural Fund administration were rather sceptical about the evaluation capacity of their respective institutions, while they were more positive about their personal evaluation capacity. (At this point it is interesting to note that evaluation capacity analysis carried out by the Ministry of Finance in 2007 demonstrates that civil servants are somewhat less positive about their personal evaluation capacity than they were in 2005. One of the possible explanations for this, as noted by one of the interviewees, is that people have already gained experience in evaluation and became more critical or realistic about their knowledge in this field.)

Once the training needs assessment had been conducted, training programme on evaluation was developed and 39-hour training for civil servants and external experts was held. To ensure the continuity of training after the end of project, 4 evaluation lecturers were trained. In addition, a shorter version of the same training programme was accredited and included into a list of training courses offered to the civil servants. Since then, training according to the programme on Structural Fund evaluation has been held twice by the Training Centre of the Ministry of Finance.⁸⁴

Another activity of the evaluation capacity building project was drafting of the methodological guidance documents: (i) the evaluation guide to assist in planning, organisation and utilisation of Structural Fund evaluation in Lithuania, and (ii) the methodological guidelines on indicators for monitoring and evaluation. *Guidelines for Evaluating EU Programmes in Lithuania* contain a description of requirements of the EU Structural Funds Regulations for the periods 2004-2006 and 2007-2013, overview the main evaluation issues and methods, and present the evaluation terminology.⁸⁵ *Guidelines for Monitoring EU Programmes in Lithuania* outline EU regulatory requirements and guidance in respect of the monitoring aspects, provide guidance on the selection of monitoring indicators, address methodological issues in quantifying the outcomes of Structural Funds interventions, and highlight good practices in managing the monitoring function. Interview evidence suggests that developed guidelines were referred to by institutions and were considered

⁸⁴ Information provided by the Training Centre of the Ministry of Finance.

⁸⁵ Guidelines for Evaluating EU programmes in Lithuania. – 2005.

quite helpful. In view of this, Evaluation Unit updated evaluation guidelines in order to reflect the most recent developments.

Finally, awareness raising measures were implemented. Of these, a regional conference on the evaluation of the EU structural support is worth mentioning. The conference entitled “Development of Evaluation Capacity and Formation of Evaluation Culture in the Baltic States” was held in November 2005. During the conference the evaluation capacity in the Baltic States was discussed and good practice shared with representatives of the European Commission and other Member States. The total number of participants amounted to 88.

All in all, the feedback provided by project participants and interviews carried out for this case study leave no doubt that evaluation strengthening project was a timely, well-received and beneficial initiative.

Source: Project material.

In spite of evaluation capacity building measures undertaken to date and certain practice accumulated during the 2004-2006 period, a need for further training in Structural Fund evaluation remains. This is evidenced by recent results of the evaluation capacity survey carried out by the Ministry of Finance in 2007. Only the general understanding about the importance and benefits of evaluation was assessed as good by the survey participants, while knowledge of specific evaluation issues was assessed as weak. Institutions still face difficulties in developing high-quality terms for evaluation projects and in quality control. In addition, they feel a lack of information on organisation of evaluation function in other EU Member States.

Training options, especially at an advanced level, are limited in Lithuania. Apart from the evaluation training course developed in the framework of the evaluation capacity building project mentioned earlier, none of the training programmes offered to civil servants deal exclusively with evaluation of the Structural Fund programmes. The Lithuanian Institute of Public Administration, which is responsible for the training of civil servants, organises a course on Strategic Management of the Structural Fund Programmes where some attention is paid to evaluation issues but evaluation is just one of the issues discussed. As a result, civil servants responsible for evaluation of Structural Fund programmes go for training courses abroad (e.g., at the European Institute of Public Administration in Maastricht).

The number of events concerning evaluation of the Structural Fund programmes is also limited. One of the most recent evaluation events was an open meeting of the Monitoring Committee which took place in May 2007. During this meeting Structural Fund evaluation experience in 2004-2006 and arrangements for the new period were presented. External evaluators were invited to participate in this meeting.

Efficiency of Evaluation Management

In the 2004-2006 period an evaluation plan was developed by Managing Authority on the basis of information provided by other institutions on evaluations foreseen, their budget and timing. The plan was discussed by the EMG. The document had no official status given the absence of regulatory requirements with regard to planning.

A new national Evaluation Plan covering the period of 2007-2013 is currently in preparation. It is being drafted by the Evaluation Unit of the National Coordinator in close cooperation with other institutional stakeholders to ensure that the plan reflects evaluation needs. The National Evaluation Plan will be approved by the Monitoring Committee once it is agreed on. At the time of writing, the plan is not available publicly as it is not finalised yet. According to preliminary information, the National Evaluation Plan will review evaluation experience in 2004-2006 programming period, set out requirements stemming from the legislation in force, and present national evaluation needs, objectives and evaluation management scheme.

Annual evaluation plans will be devised for the implementation of the multi-annual evaluation plan. These will set out in detail specific evaluation and evaluation capacity building projects, timing, project budget and responsible institutions. Evaluation projects will be included in the annual evaluation plan if they conform to certain criteria, a number of which will derive from Structural Fund Regulations.

It is expected that the planning of evaluation activity will assist in coordination of evaluation function and contribute to better quality evaluations. This will require effective implementation and supervision of evaluation plans. The current thinking is that one measure that would promote effective planning might be a rule whereby financing would be made available only to evaluation projects included in the plan. Such a requirement would reflect lessons of the 2004-2006 programming period when institutions could submit applications for technical assistance for evaluation without informing the MA or the EMG, a practice which did not support coordination of evaluation.

At the time of writing, it has not yet been decided what would be the procedure for handling ad hoc evaluations which are not planned and included in annual evaluation plan, but the need for which arises in the course of programme implementation (e.g. when certain difficulties in implementation arise). The current thinking is that several measures could be applied, such as certain periodic revision of annual evaluation plans (where external evaluation services are needed) or internal evaluations carried out by the Evaluation Unit, assisted by a consultant for evaluation (for smaller scale evaluations).

With reference to other stages of the evaluation management process, it can be repeated that national authorities, in particular intermediate bodies, still face difficulties in developing terms of references for evaluation projects. The quality of terms of reference varies across institutions. A general tendency is to develop more specific and focused terms of reference. To some extent this reflects the reality that institutions become more definite about their needs and also attempt to use terms of reference as a quality control tool. While high degree of standardisation of terms of references for evaluation assignments is not practicable, some standardisation might be worth considering (e.g. having core terms of references for strategic and operational evaluations).⁸⁶

As regards the time needed to prepare terms of reference, it takes approximately 2-3 months to draft and agree on terms of references and tendering requirements. The length of the subsequent procurement process depends on the procedure chosen. In case of an international⁸⁷ open procedure, the procurement process usually takes four to five months from prior information notice to the conclusion of a contract. The deadline for receipt of tenders cannot be less than 52 days from the date of dispatch of the contract notice from the Public Procurement Office. Under the restricted procedure a time limit for submission of requests to participate cannot be less than 37 days from the date of dispatch of the notice from the Public Procurement Office.⁸⁸ Under the negotiated procedure there is no requirement to publish a contract notice. Under this procurement mechanism, the contracting authority negotiates contract terms and conditions directly with one or several invited suppliers. Finally, under the negotiated procedure with publication of a contract notice a time limit for submission of requests to participate cannot be less than 37 days from the day of dispatch of the notice from the Public Procurement Office. The time limit for receipt of tenders may not be shorter than 40 days from the day of dispatch of the invitation to tender.⁸⁹

Service contracts may be awarded on the basis of one of two criteria: either the lowest price or the most economically advantageous tender. As a rule the latter criterion is applied. This is also

⁸⁶ Observation made by one of the interviewees.

⁸⁷ International procedure has to be organised when the value of evaluation services exceeds 154 000 EUR. Simplified procedure can be applied when the value of evaluation services is less than 154 000 EUR. Usual commercial practice is applied when value of evaluation services does not exceed 21 721 EUR.

⁸⁸ Law on the Public Procurement, 13 August 1996, No. I-1491. - New version 22 December 2005, No. X-471

⁸⁹ Guidelines for Evaluating EU programmes in Lithuania. – 2005.

recommended by the *Guidelines for Evaluating EU Programmes in Lithuania*. On average the price-quality ratio is of 60:40 or 40:60.⁹⁰

For larger evaluation projects (e.g. ex ante evaluation of OPs for 2007-2013) project steering groups are usually set up. As pointed out by the authorities, steering and quality control are one of the most demanding tasks. Measures, which are considered to be the most important in ensuring quality control, are good quality of terms of references, competition between tenderers and bigger attention to the quality of inception reports in order to make sure that appropriate methodology to evaluation tasks is adopted in advance.

Quality of Monitoring Systems

Both administrators and external evaluators acknowledge that there is still room for improvement as regards precision, quality and timeliness of monitoring data, especially physical data. Concerns were expressed regarding the definition of indicators and timeliness of inputting data into the system, etc. Survey indicates that the average interval between year end and availability of monitoring data is approximately 3-6 months, but the timeliness also depends on the type of indicator. It is worth noting that an evaluation aimed at improving Structural Fund monitoring and information system and adapting it to requirements of 2007-2013 programming period, is being planned.

(4) Supply Side

Having discussed the demand side, the analysis below focuses on supply-side criteria included under the supply dimension of the Evaluation Capacity Development Benchmarking Framework, namely:

- Quality of socio-economic data resources; and
- Availability and quality of evaluation expertise.

Quality of Socio-economic Data Sources

Availability of reliable external data on key socio-economic variables is essential for evaluation of Structural Fund programmes with a wide range of policy objectives. In general, the quality and timeliness of contextual, socio-economic data (e.g., GDP, employment, unemployment, R&D investment) at national level is felt to be adequate in Lithuania. The latest year for which most data is available is 2006.

While the key statistical indicators are also available at the county (regional) level, more specific sectoral data needs to be collected from various specialised institutions (e.g. data on telecommunications - from the Communications Regulatory Authority, etc) and is not always to hand.

Availability and Quality of Evaluation Expertise

To date Lithuanian authorities have relied on external services for the evaluation of Structural Fund programmes. Institutions do not carry out internal evaluations due to insufficient staff resources and expertise. However, as already mentioned above, there are intentions to start doing some small-scale evaluations at the Evaluation Unit.

Demand for evaluations in the 2004-2006 period stimulated the emergence of small evaluation market in Lithuania. It is still dominated by several consulting companies. Some 10 companies could be listed as offering evaluation services, but as a rule 4-5 service providers compete in the tendering process. Authorities feel that this degree of competition, especially when it comes to the quality of tenders, is not sufficient yet. Authorities note that quality of tenders is variable, with just

⁹⁰ Public Procurement Office recommends that price weight is no less than 40 percent.

a few companies submitting higher quality tenders. As a result the same few companies carry out most of the Structural Fund evaluation assignments and, as a result, it is difficult for newcomers to accumulate experience and references and to enter the market. With the exception of the agricultural sector, there is little sectoral specialisation with the same companies tending to participate in calls for tenders announced by different institutions. This is not surprising as during short programming period of 2004-2006 the demand for evaluations has not been such as to allow for specialisation in evaluations in certain sectors.

As regards the type of service providers, a few foreign companies are active in the evaluation market in Lithuania. While the number is small, they have comparatively good positions. The tendency is for consortiums between domestic and foreign service providers to form to carry out specific evaluation projects (e.g. ex ante evaluation of the SPD and ex ante evaluation of OPs for 2007-2013, also the evaluation capacity building project). Universities or research institutions are virtually absent from the evaluation market, except the agricultural sector. Then again, a small number of researchers are involved in evaluation via consulting companies, in particular for assignments requiring macro-economic modelling. There are intentions from the authorities to activate academic community and attract its attention to Structural Fund evaluation, for example, by encouraging students to write their theses in the area of Structural Fund evaluations.

The choice of education and training in evaluation of Structural Fund programmes for external evaluators is limited in Lithuania. Thus inviting evaluators to participate in several evaluation capacity building events carried out so far was a welcome initiative. External experts may attend the same courses offered to civil servants by the Lithuanian Institute of Public Administration or the Training Centre of the Ministry of Finance, but these are mostly “beginner’s” courses. Besides, little attention is paid to such issues as evaluation methods or evaluation theory.

(5) Dissemination and Utilisation

This section considers issues related to the publicity and utilisation of evaluation results, which fall under the third dimension of the evaluation capacity development benchmarking framework.

Information on evaluation of the Structural Fund programmes and all evaluation reports commissioned by the MA are published on the webpage <http://www.esparama.lt> dedicated to EU structural assistance issues. Some evaluations contracted by intermediate bodies are available on their institutional websites. There are intentions to create a centralised online access to all Structural Fund evaluation reports in Lithuania. Whether this will be implemented remains to be seen.

Even though it is not yet a widespread practice, there are instances where Structural Fund evaluation findings have been discussed quite widely with a number of stakeholders. A good practice example is a public discussion of findings of the evaluation of the Structural Funds administration in education and science sector, organised by the Ministry of Education and Science (intermediate body). Another good practice example is dissemination of evaluation findings concerning the Structural Fund publicity and information system - findings of this evaluations were presented and discussed by institutions and social-economic partners in a special seminar.

As regards utilisation of evaluation results, there are no institutionalised procedures for addressing evaluation findings or recommendations. Evaluation results are discussed at the EMG, evaluation steering group and the Monitoring Committee. Certain experience with formalised procedures for dealing with evaluation results was accumulated when implementing Phare.⁹¹ However, this experience shows that utilisation and tracking of evaluation recommendations was done in a somewhat formal manner.⁹²

⁹¹ Monitoring reports of PHARE projects indicate actions taken to implement evaluation recommendations. Moreover, responsible institutions have to report to the Ministry of Finance for the implementation of recommendations. // Nakrosis, Vilpisauskas, 2003. p. 22.

⁹² Nakrosis, Vilpisauskas, 2003. p. 22.

At the time of writing of this report it is not clear whether any institutionalised procedures for addressing evaluation results will be developed for the period of 2007-2013. However, a general feeling is that such measures might have adverse effects, e.g. discouraging the usage of evaluation and making it more of a control tool rather than a learning experience.

In terms of utilisation of the Structural Fund evaluations, external evaluators note that some recommendations have been implemented. Evaluation timing and quality of recommendations, among other factors, influence the extent to which evaluation results are used. The authorities acknowledge that the timing was not always well planned so as to allow for utilisation of evaluation results in 2004-2006 period.⁹³ In view of this experience, evaluation plans are expected to address such challenges.

As regards impact of the recent ex ante evaluation, the survey and interviews suggest that limited influence was made on the distribution of resources and overall strategy, while the impact on quantitative targets and situation analysis was significant. The final ex ante evaluation report put forward 27 general and more than 180 specific recommendations concerning individual OPs. During the programming process more than two-thirds of evaluators' recommendations were utilised.⁹⁴

(6) Degree of Institutionalisation of Evaluation in the Wider Governance System

Finally, this section considers various criteria specified under the institutionalisation dimension of the Evaluation Capacity Development Benchmarking Framework as follows:

- The role of evaluation in the national budgetary process;
- The role of evaluation in the formulation of sectoral strategy and policy making;
- The existence of mandatory evaluation requirements;
- The public policy research base;
- National evaluation society;
- Influence of EU evaluation requirements; and
- The role of civil society.

The national evaluation system in Lithuania consists of impact assessments of the Government's draft decisions and evaluation of budgetary programmes. These evaluation measures are regulated by Lithuanian legislation. Impact assessments of the Government's decisions are carried out according to the Government's resolution of 2003.⁹⁵ According to the respective resolution, a basic or extended⁹⁶ impact assessment of a draft decision is conducted.⁹⁷ Impact assessments still face certain problems: results of basic evaluations are rather general while extended assessment is applied rarely. The quality of impact assessments varies and utilisation of assessment results is rather limited.⁹⁸

The evaluation of budgetary programmes is also performed according to the specific Government's resolution.⁹⁹ Since 2000 appropriation managers have been preparing and updating strategic plans and budget programmes, the implementation of which are monitored and evaluated. The evaluation of national budgetary programmes is first of all a result of a strategic planning reform in Lithuania.

⁹³ Stankaitiene A. Evaluation Capacity Building in Lithuania. – Presentation for Evaluation Units Open Days. – 27 June 2006, Italy.

⁹⁴ Information provided by the Evaluation Unit.

⁹⁵ The Resolution of the Government of the Republic of Lithuania No.276 of 26 February 2003 “On Approval and Implementation of the Methodology for Assessing Impact of Draft Decisions”.

⁹⁶ Extended assessment should indicate monitoring and accountability arrangements. However, the extended evaluation is rarely applied. // Nakrosis, Vilpisauskas, 2003. - p. 20

⁹⁷ Nakrosis, Vilpisauskas, 2003. - p. 18

⁹⁸ Nakrošis and Vilpišauskas, 2003 - P. 19-20.

⁹⁹ The Resolution of the Government of the Republic of Lithuania No.194 of 7 February 2007 “On Methodology for Strategic Planning”.

Evaluation of budgetary programmes is internal in nature: it is carried out by internal audit units. Ex ante evaluation of budgetary programmes is not conducted. The only evaluation criterion assessed is effectiveness. Other evaluation criteria such as relevance or efficiency are not applied. Another concern is that evaluation recommendations are not sufficiently used in the decision making.¹⁰⁰

In addition to the measures mentioned above, Lithuanian ministries have available financing for various research studies, services for which are purchased externally. The Prime Minister's Office has annual financing for analytical studies on impact of Lithuania's integration into the EU.

Regardless of existent national evaluation measures, it must be acknowledged that a tradition of evaluation and orientation to results in Lithuanian public sector is limited. A demand for evaluation was driven mostly by external regulatory requirements and only partially by the administrative reform.¹⁰¹

As to the supply side, there are a few organisations involved in public policy research (e.g. Lithuanian Free Market Institute, Civil Society Institute, Institute of International Relations and Political Science of Vilnius University and a few consultancy companies). A few of them are also interested in Structural Fund related issues.

There is no professional organisation of evaluators in Lithuania. Views as to the need for and use of such an organisation are mixed.

Finally, with reference to a wider civil society, a certain interest in results of Structural Fund evaluation exists. There has been some debate on issues concerning Structural Fund institutional system and investment areas.

(7) Conclusions

In sum, the demand and supply sides of Structural Fund evaluation are developing in Lithuania where a limited tradition of evaluation existed prior to EU accession. Demand for Structural Fund evaluation was primarily driven by external requirements, but in 2004-2006 programming period these concerned only the ex ante evaluation of the SPD. However, national authorities decided to launch a number of ad hoc evaluations. Most of their results are publicly available and have attracted some attention from a wider public.

In view of practically non-existent evaluation capacity, authorities made efforts to implement measures aimed at building evaluation capacity both within the administration and externally. Further evaluation strengthening activities are being planned for 2007-2013 programming period and will be included in the National Evaluation Plan.

Both in 2004-2006 and 2007-2013 programming periods Lithuania took a centralised approach to Structural Funds evaluation management. In 2006 a separate Evaluation Unit was established in the National Coordinator to coordinate evaluation activity at national and OP level. This is an institution which has built up the most extensive evaluation management experience in the Lithuanian context.

There are still several obstacles to evaluation capacity development within the wider Structural Fund administration in Lithuania. The number of full-time evaluation staff in Intermediate Bodies appears to be less than sufficient. Shortcomings in regulation of responsibility for evaluation function also act as a constraint on the development of administrative evaluation capacities.¹⁰² However, there are already plans at some institutions to address the issue of human resources, not least due to the increasing need for evaluations.

¹⁰⁰ Nakrosis and Vilpisauskas, 2003. - p. 13-14; 20; 78.

¹⁰¹ Nakrošis and Vilpišauskas, 2003

¹⁰² Evaluation capacity analysis carried out by Ministry of Finance, April-May 2007

A challenge facing authorities in the 2007-2013 programming period is related to on-going evaluation, as there is no experience with such an approach to evaluation in Lithuania. It will be important that authorities carefully estimate and plan their evaluation needs, which requires certain evaluation capacity.¹⁰³ The quality of the monitoring system and data is also an important issue.

As regards the supply side, a small market of evaluation services emerged in 2004-2006 programming period as institutions relied exclusively on external evaluation services. The market is still dominated by a few consulting companies and needs further development, both in number and type of service providers as well as the quality of services. Systematic planning of evaluation demand and more evaluations commissioned might potentially serve as a catalyst for the development in 2007-2013 programming period.

(8) Key Learning Points

With regard to evaluation capacity benchmarking framework

- As regards the application of the Evaluation Capacity Benchmarking Framework to Lithuania case study, it faced practically no major difficulties. All criteria and indicators could be assessed.

From Lithuanian experience

- As regards Lithuania's experience in Structural Fund evaluation, one of the key learning points is evidently attention paid to evaluation capacity building. Lithuania addressed the issue of evaluation capacity at the initial stage of its development and at both the demand and supply sides of the market. As preparations for the new period indicate, the authorities intend to continue with capacity building initiatives.
- While it is too early to assess the organisation of evaluation function in 2007-2013 programming period as it is still under development, it is worth noting that steps are being taken to aim at better planning, monitoring and coordination of evaluation activity.

¹⁰³ As noted by one of the interviewees.

Bibliography

1. Evaluation capacity analysis, carried out by the Ministry of Finance in April-May 2007. Presentation made on 25 May 2007 during the open Monitoring Committee Meeting.
2. Evaluation of the implementation system of Lithuanian Single Programming Document 2004-2006
http://www.esparama.lt/lt/bpd/administravimo_sistema/BPD_igyvendinimo_vertinimas
3. Guidelines for Evaluating EU programmes in Lithuania. – 2005.
4. Law on the Public Procurement. 13 August 1996, No. I-1491. - New version 22 December 2005, No. X-471. – Available in English at
<http://www.vpt.lt/admin/uploaded/LawonPPnew.doc>
5. Lithuania. Cohesion Policy 2007-2013. // http://ec.europa.eu/regional_policy/atlas2007/fiche/lt_en.pdf
6. Order of the Minister of Finance of 30 May 2005 No. 1K-175 „On Composition of the Evaluation Management Group of Lithuanian Single Programming Document for 2004-2006“.
7. Presentations made at the regional conference „*Building evaluation capacity and culture in the Baltic States*“, organised in Vilnius, 28-29 November 2005.
8. Project material „Strengthening Evaluation Capacities“ (carried out June 2005 – February 2006. Contracting authority Ministry of Finance.)
9. Stankaitienė A. Evaluation Capacity Building in Lithuania. – Presentation for Evaluation Units Open Days. – 27 June 2006, Italy.
10. The Resolution of the Government of the Republic of Lithuania No. 1166 of 11 September 2003 “On the Administration of Assistance from the EU Structural Funds in Lithuania”.
11. The Resolution of the Government of the Republic of Lithuania No.276 of 26 February 2003 “On Approval and Implementation of the Methodology for Assessing Impact of Draft Decisions”.
12. The Resolution of the Government of the Republic of Lithuania No.194 of 7 February 2007 “On Methodology for Strategic Planning”.
13. Vilpišauskas R., V. Nakrošis “Viešosios politikos vertinimas Lietuvoje“, Vilnius, 2005. (eng. „Public Policy Evaluation in Lithuania“)

Annex to Lithuania case report: List of Structural Fund evaluations commissioned in Lithuania during 2003-2007

No.	Contracting Institution	Evaluation project	Period
1.	Ministry of Finance (Managing Authority)	„Ex ante evaluation of Single Programming Document for 2004-2006“	2002 – Oct 2003
		„Thematic evaluation of the SPD and evaluation capacity building“ (consisting of two reports): (i) Evaluation of effectiveness of the SPD implementation system for 2004-2006 (ii) Structural Fund investment areas for 2007-2013 (http://www.esparama.lt/lt/bpd/administravimo_sistema/BPD_igyvendinimo_vertinimas)	April 2005– May 2006
		„Planning optimal and efficient structure of the EU Structural Fund programmes and support management institutional system“	May 2005 – December 2006
		„Ex ante evaluation of the OPs for 2007-2013“ (available in English and Lithuanian at http://www.esparama.lt/lt/pasirengimas/vertinimas)	February 2006– June 2007
		Evaluation of implementation of horizontal priorities	2007
		Evaluation of the communication and information about the EU Structural Funds	2007
2.	Ministry of Education and Science (Intermediate Body)	„Lessons from 2004-2006 programming period and challenges for the EU Structural Fund support for science and education projects in 2007-2013“	March 2006– May 2007
3.	Ministry of Transport and Communications (Intermediate Body)	„Evaluation of Structural Funds investment in transport sector and recommendations for the 2007-2013 programming period“	Nov 2006 – July 2007
4.	Ministry of Agriculture (Intermediate Body)	„The match between the chosen Structural Fund support measures and their implementation and society’s expectations“ (ppt presentation http://www.zum.lt/min/failai/ZUM_tyrimo_prezentacija_2007-02-06.ppt)	June 2006 – Nov 2006
		“Effectiveness of the consulting system”	Aug 2006 – Feb 2007
		“Effectiveness of the beneficiaries’ participation under the SPD measure “Investment in agricultural holdings“	Aug 2006 – Feb 2007
		“Consequences of the farmers’ lifelong learning supported by Structural Funds for entrepreneurship”	Aug 2006 – Nov 2006
5.	Ministry of Social Security and Labour (Intermediate Body)	Analysis of human resource sector: situation, prospects and recommendations on the investment strategy in the 2007-2013 Operational Programmes in the areas, which fall under the responsibility of the ministry.	July 2006 – December 2006
6.	Ministry of Interior	„Study of by Structural Fund investment in public administration, local and regional development in 2007-2013“ http://www.vrm.lt/index.php?id=764	April 2006 – August 2006
7.	Information Society Development Committee under the Government (Intermediate Body)	Study on Structural Fund investment in the Information Society Development field in 2007-2013	May 2006 – Sept 2007
8.	National Paying Agency (Implementing Body)	„Study on potential risks for the absorption of SPD support for agriculture“	April 2006 – May 2006

		„Study on absorption of SPD support for agriculture”	July 2006 – August 2006
9.	Central Project Management Agency (Implementing Body)	Beneficiary attitudes towards training, preparation of applications and project implementation and assessment of information needs	July 2005 – October 2005

Report on Romania Case Study

(1) Introduction and Methodology

This report presents the results of the case study of evaluation capacity in Romania. The case study is based on a review of relevant documentation and a series of semi-structured interviews conducted in Bucharest on 28 and 29 May last.

Some 7 interviews were conducted in Bucharest involving the participation of some 12 persons in total. The interviewees comprised representatives of three central government institutions, the European Commission Delegation to Romania and three private sector evaluation practitioners - one of whom acts as coordinator of *Evalrom*, the informal Romanian evaluation network. At central government level, the interviews were held with representatives of the General Secretariat of the Government, the Ministry of Economy and Finance (overall authority for coordination of EU Funds in Romania) and the Ministry of Interior and Administrative Reform (which will act as managing authority for the Administrative Capacity Development Operational Programme 2007-2013). The full list of documents reviewed is set out in the annexed bibliography.

The documentary analysis and the interviews provided a wealth of useful information on the various components (i.e., dimensions, criteria, sub-criteria and indicators) of the evaluation capacity development benchmarking framework. In some cases, hard quantitative data was not readily to hand (e.g., indicator number 6 relating to evaluation budgets, no. 12 relating to procurement intervals and no. 19 concerned with the number of firms active in the evaluation market) although some useful impressions or views on the issues in question were generally conveyed.

(2) Background Information: Structural Funds in England

As Romania acceded to the European Union in 2007, its first experience of the Structural and Cohesion Funds (SCF) will be in the 2007-2013 period. However, Romania has been in receipt of pre-accession funds for some time and reference will be made to the experience of evaluation of the pre-accession Phare programme in a number of instances below. In particular, reference will be made to a number of Phare projects aimed at supporting the development of evaluation capacity.

For the 2007 to 2013 period all regions in Romania are eligible under the Convergence objective. Romania has been allocated approximately €19.2 billion under this objective and €455 million under the Territorial Co-operation objective. The strategic objectives and priorities of the Romanian National Strategic Reference Framework (NSRF) will be implemented through 7 operational programmes as follows:

NSRF Operational Programmes¹⁰⁴

Operational Programme	Managing Authority	Total SCF Allocation (€million)
Transport	Ministry of Transport, Construction and Tourism	4566
Environment	Ministry of Environment and Water Management	4512

¹⁰⁴ Excluding programmes under the territorial cooperation objective.

Increase of Economic Competitiveness	Ministry of Economy and Trade	2554
Regional	Ministry of European Integration	3726
Human Resources Development	Ministry of Labour, Social Solidarity and Family	3476
Administrative Capacity Development	Ministry of Administration and Interior	208
Technical Assistance	Ministry of Economy and Finance	170

(3) Demand Side

The analysis is presented below under the 4 criteria included under the demand-side dimension of the evaluation capacity development benchmarking framework. These are as follows:

- The integration, location and organisation of the Structural and Cohesion Funds (SCF) evaluation function;
- Evaluation resources;
- Efficiency of evaluation management;
- Quality of monitoring systems.

Integration, Location and Organisation of Evaluation Function

Dedicated evaluation structures with responsibility for organising the evaluation function have been put in place at both NSRF and programme level in Romania. At overall NSRF level, the Romanian authorities have established an Evaluation Central Unit (ECU) in the Managing Authority Directorate of the Ministry of Economy and Finance. For each of the operational programme listed above (with the exception of the Technical Assistance OP), evaluation units have been established at managing authority level. In addition, there are two units dealing with the evaluation of the cross-border cooperation programmes (one with programmes with EU Member States, the other with non Member States, both attached to the Ministry of European Integration). A standard text is included in each OP setting out the overall evaluation arrangements and the functions of the evaluation units at each level. The ECU is responsible for overall coordination of evaluation across the NSRF including organising cross-cutting evaluations, capacity building activities and quality assurance. The OP evaluation units are responsible for commissioning interim and ad hoc evaluations of their respective programmes. A common Procedures Manual has been developed for the units setting out in more detail their functions and working methods.

As noted above, the ECU is responsible for overall coordination of evaluation across the NSRF. Coordination is effected through the Evaluation Working Group (EWG) which is chaired by the ECU and includes representatives of each of the OP units. Since early 2006, the EWG has met on a regular basis. For much of 2006 the emphasis was on various training and capacity-building activities supported under a Phare Technical Assistance project (*Technical Assistance in Programming, Monitoring and Evaluation*, RO 2003/005-551.03.03.04) In addition, working groups have been established on a number of issues of common interest including management information systems and cost-benefit analysis.

In each case, the evaluation function is located within the Directorate in the Ministry with responsibility for the Managing Authority function. The evaluation function is often combined in a Directorate/Managing Authority unit with responsibility for a related function such as programming or monitoring.

Evaluation Resources

Information on the evaluation budget was not available in respect of each OP at the time of the case study. The budget for the ECU under the Technical Assistance OP amounts to €10.4 million of which €8.3 represents EU funding. Under the Transport OP, the amount allocated for evaluation will be in a range of between €13.8 million and €17.25 million (includes national contribution).

In terms of staffing resources, the intention is that the ECU will have up to 5 full-time staff with the OP units staffed by 2 personnel each. Not all of the units are yet operating at these levels and the recruitment and retention of skilled staff could prove problematical. Reflecting the reality that evaluation practice has been very limited in Romania, it is understandable that the personnel have very little evaluation experience although a number would have participated in some capacity in evaluations of the Phare programme. More recently, some of the staff participated in steering committees for the ex ante evaluations of the 2007-2013 programmes; this was felt to have constituted a useful learning experience.

The only evaluations carried out to date of SCF programmes were the above-mentioned ex ante evaluations. The standard OP text outlines the intention to undertake interim evaluations of each OP in 2009 and 2012 with additional ad hoc evaluations as needs dictate.

Efficiency of Evaluation Management

Working under the aegis of the Evaluation Working Group, each evaluation unit has already prepared a draft OP evaluation plan based on the guidance in the Commission's Working Document No. 5.¹⁰⁵ The draft plans are quite comprehensive in scope and contain specific proposals for evaluation projects at NSRF and OP level respectively. The plans will be finalised once agreement on the programmes has been reached and will be subject to the approval of the relevant monitoring committee.

The benchmarking indicators relating to terms of reference (no. 11) and procurement intervals (no. 12) cannot be assessed at this stage given the very limited SCF evaluation experience. The terms of reference for the ex ante evaluation followed very closely the key components listed in the Commission's Working Document No. 1.¹⁰⁶

As regards evaluation steering arrangements, each OP will have an Evaluation Steering Committee to be convened for each evaluation project. At NSRF level, a Strategic Evaluation Steering Committee will be established for cross-cutting evaluations. These committees will fulfill the standard steering functions including: setting the terms of reference for individual evaluations; facilitating and supporting the evaluator in his/her work particularly as regards methodological issues; ensuring that the terms of reference are observed and exercising quality control. Quality assurance for the ex ante evaluations was carried out on the basis of the quality criteria included in the Guide to the Evaluation of Socio-Economic Development (EVALSED)¹⁰⁷ and referred to in the Commission's Working Document No. 1 (the quality matrix is also included in the Procedures Manual referred to earlier).

¹⁰⁵ European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: Methodological Working Papers. *Indicative guidelines on evaluation methods: evaluation during the programming period*. Working Document No.5.

¹⁰⁶ European Commission, Directorate-General for Regional Policy. The New Programming Period, 2007 – 2013: *Indicative Guidelines on Evaluation Methods: Ex Ante Evaluation*, Working Document No.1

¹⁰⁷ EVALSED: the on-line resource for the evaluation of socio-economic development

Quality of Monitoring Systems

The quality of the OP monitoring systems – in terms of the quality, reliability and timeliness of indicator data - cannot be assessed at this stage given that programme implementation has not commenced. Some concerns were expressed as regards the reliability of Phare programme monitoring data, responsibility for which has rested with the Romanian authorities since 2002.

(4) Supply Side

The analysis below of the supply-side is presented under the two criteria included under the supply dimension of the evaluation capacity development benchmarking framework, namely:

- Quality of socio-economic data resources; and
- Availability and quality of evaluation expertise.

Quality of Socio-Economic Data Resources

View on the quality and timeliness of contextual socio-economic indicator data (GDP, employment, unemployment, R&D investment) were somewhat divided amongst the three external evaluators consulted. Overall, it was felt that the quality of data produced by the National Institute of Statistics was improving due, inter alia, to pressure from Eurostat. However, some concerns were expressed about the range of data available online and the limited data available at the county level.

Availability and Quality of Evaluation Expertise

The number of firms or organisations involved in the evaluation business was felt by most commentators to be rather limited (reflecting very limited demand in the past) but to be experiencing significant growth at present. Very few – if any – firms or freelance operators specialise solely in evaluation. The market was viewed as being in its infancy with a scattered expertise residing in individuals rather than institutions. The university sector has not been active in the evaluation market thus far. The question of internal evaluation unit capacity was dealt with above under evaluation resources.

Most prior evaluation experience relates to participation as local experts in the evaluation of pre-accession instruments (particularly the Phare programme). These evaluation exercises were generally managed by foreign contractors. The ex ante evaluation process for the 2007-2013 programmes was similarly organised (it was funded by Phare) but again some local experts were involved.

A number of significant capacity development initiatives have been (or are being) supported under the Phare programme. Under an earlier Phare Technical Assistance project (*Technical Assistance in Programming, Monitoring and Evaluation*, RO 2003/005-551.03.03.04), some 34 participants from the private, academic and civil society sectors graduated from a 3-month certificate course in Policy and Planning (Evaluation), accredited by the National University of Ireland, Galway in June 2006. This project also supported some six evaluation awareness-raising seminars over the period May/June 2006 (298 participants) as well as supporting the work of the ECU and the EWG referred to earlier (including in the preparation of Evaluation Plans for the 2007-2013 programmes).

A second recently-commenced *Evaluation Facility* project aims to encourage public and local bodies to introduce evaluation into their management processes. Total funding of some €800,000 will fund projects (comprising a mix of project, programme and policy-level ex

ante, interim, ex post and thematic evaluations) at central government, county and local (municipality) level. The evaluations will be carried out by external teams of international experts working alongside local experts. The project is managed by a consultancy team operating under the guidance of the ECU.

Finally, the procurement process for a Phare 2005 approved project entitled *Development of a Professional Evaluation Community* will be launched shortly. According to the *National Evaluation Strategy* (discussed below), this will aim at developing evaluation skills and standards on the supply side of the evaluation community, while also developing the normative acts necessary to be passed by the Romanian Parliament, in order to embed evaluation in the management of all public interventions, and also to develop evaluation standards and principles. Inter alia, the project is to support evaluation training (including curriculum development) and evaluation networks in Romania.¹⁰⁸

It is too early at this stage to comment on the breadth and quality of expertise across different sectors or themes. The issue at present is one of an overall or global expertise deficiency.

Finally, it is not possible at this stage to comment on the quality of SCF evaluation reports other than as regards the recent ex ante evaluation exercise. Of the seven OP ex ante reports reviewed for the ECU by external quality assessors (carried out on the basis of the EVALSED quality assessment framework included in the Commission's Working Document No. 1 referred to earlier), five were rated as satisfactory on the *Meeting Needs*¹⁰⁹ criterion, one was rated as "satisfactory to good" and one was rated as "good". For the four OP evaluations where overall ratings were recorded, 1 was rated as "satisfactory/good", 2 as "satisfactory" and 1 as "less than satisfactory".¹¹⁰

(5) Dissemination and Utilisation

As regards public access to evaluation reports, the programmes include a commitment to the effect that, at a minimum, the executive summary of evaluation reports will be made publicly available.

The OPs also indicate that, under the coordination of the ECU, a mechanism for addressing and following-up evaluation recommendations will be put in place under the responsibility of the Managing Authorities. These will be based on the processes outlined in the Procedures Manual referred to above. Based on previous Phare practice, a follow-up procedure was applied in the case of the recent 2007-2013 OP ex ante evaluations with two de-briefing meetings per programme held between the relevant evaluators and managing authority.

In terms of the impacts of evaluation on programme design and implementation, this can only be assessed in the SCF sphere in respect of the ex ante exercise. Previous experience with the Phare programme was somewhat mixed; one observer felt that the Romanian authorities had a good track record as regards implementation of what might be regarded as "technical" recommendations that fell within the competence of the Programme Implementation Unit or Ministry in question. However, there was no mechanism to take forward broader, policy-related recommendations to the appropriate higher level.

¹⁰⁸ Separately, it is understood that the University of Cluj is launching a Master's level degree course in Public Administration in October 2007 which will include a dedicated evaluation module.

¹⁰⁹ The "Meeting Needs" criterion relates to whether an evaluation report adequately addresses the requests for information formulated by the commissioners and corresponds to the terms of reference.

¹¹⁰ The author of this case study carried out the quality assessments of two OP evaluations for the ECU.

As regards experience with the ex ante evaluation exercise, it was felt that the reports had little influence on the allocation of resources as decisions on these issues were already taken by the time the evaluations commenced. On the other hand, the evaluators helped the managing authorities in the articulation of the SWOT analyses and programme strategies. They also provided lists of possible indicators which contributed to improvements in this area in some instances. Overall, it was felt that the exercise had some positive influence on the programming process.

(6) Degree of Institutionalisation of Evaluation in the Wider Governance System

The section below considers the following five criteria specified under the institutionalisation dimension of the evaluation capacity development benchmarking framework:

- The role of evaluation in the national budgetary process;
- Role of evaluation in the formulation of sectoral strategy and policy making;
- The existence of mandatory evaluation requirements;
- Public management reforms; and
- The public policy research base.

The criteria relating to the existence of a national evaluation society, the influence of EU evaluation requirements and the role of civil society are considered later.

For historical reasons, evaluation and related forms of analysis have not featured strongly in the public policy making process in Romania. There is some tradition within ministries in the production of action plan monitoring reports but these have tended to be of a “tick-box” character and to focus unduly on positive developments rather than giving a rounded, objective view of performance. Major problems identified with the policy making process include the reality that the process has traditionally been dominated by a legislative approach with poor linkage to the budgetary process, a rather reduced role for civil servants and insufficient skills in areas such as impact assessment and financial analysis.¹¹¹

However, under the umbrella of the Public Administration Reform (PAR) strategy that the Romanian Government has pursued since 2001, a number of significant reforms to the policy making process have been put in train in recent years. These reforms have been driven domestically by the General Secretariat to the Government (GSG) with external pressure and support from the World Bank, the International Monetary Fund and the European Commission. Support from the Phare programme has been of particular significance; as noted in the proposed Administrative Capacity OP, Phare 2002 to 2004 commitments of over €72 million have supported PAR projects in a number of priority areas including in the Ministry of Economy and Finance, the GSG, the Court of Account and the Parliament.¹¹²

A number of the more significant recent initiatives or “milestones” include the following:

- The establishment of a Public Policy Unit in the GSG in 2003 and subsequent establishment of counterpart units in the line ministries staffed with young, qualified analysts and aimed at enhancing ex ante policy analysis capacity;

¹¹¹ *Strategy for improving the public policy making, coordination and planning system at the level of the central public administration*, General Secretariat to the Government, 2006

¹¹² See the Ministry of Interior and Administrative Reform revised Official Administrative Capacity Development Proposal dated May 2007.

- The adoption of Government Decisions 775 and 1361 which imposed requirements on ministries regarding the procedures and formats (including ex ante impact assessment) to be followed in developing public policy and normative act proposals for consideration by Government;
- The development by the GSG (with assistance from the State Chancellery of Latvia under the aegis of a Phare project) and adoption by Government (decision 870/2006) of the *Strategy for improving the public policy making, coordination and planning system at the level of the central public administration*;
- Adoption of Government decision 1807 of 2006 which requires ministries to introduce a system of strategic planning with the aim of bringing together policy planning, budgeting, priority setting and organisational framework in one management framework. In a first phase, ministries produced a management component of these plans covering the period 2007-2009 (drawing on a guidance document issued by the GSG). In a later second phase, it is intended to link this process with the budget process in parallel with the development by the Ministry of Economy and Finance of a medium-term expenditure framework. Work is also now underway in the GSG on the development of policy monitoring and evaluation methodologies and systems.
- The launch by the Ministry of Economy and Finance in November 2006 of the *National Evaluation Strategy 2007-2013*. In general terms, the Strategy outlines a pathway intended to lead to a common, integrated approach to evaluation of all public interventions in Romania, whether funded by the EU or the national budget by 2013.¹¹³ The intention is that the implementation of the Strategy will be supported by a number of funding sources including the Phare *Evaluation Facility* and *Development of a Professional Evaluation Community* projects (referred to above), the Administrative Capacity OP for 2007 to 2013 and technical assistance funding for evaluation under the other OPs.

Looking ahead, the Administrative Capacity OP (ESF budget of €208 million) will be a key vehicle for supporting and indeed joining-up these initiatives and reforms. In particular, Priority Axis 1 *Structural and Process Improvements to the Public Policy Management Cycle* will support the following key areas of intervention:¹¹⁴

- Improve political-administrative decision making;
- Strengthen the accountability framework, which is to focus on the “implementation (reporting) and *evaluation* elements of the public policy management cycle” (emphasis added); and
- Improve organisational effectiveness.

It is envisaged that the GSG and other government ministries will be amongst the beneficiaries of training and other activities that will be supported under this Priority Axis.

In summary, while Romania’s baseline position in terms of these benchmarking criteria/indicators (relating to the role of evaluation in the budget and policy making processes, existence of statutory evaluation requirements and a public policy research base) might be characterised as relatively low by international standards, a significant number of reform

¹¹³ The overall aim of the Strategy is that “there will be a functioning national evaluation system, the parts of which reinforce each other; encompassing the public and private sector and civil society; and contributing to the effective management of public interventions and the accountability of policy makers and public managers. There shall be recognition of the importance of evidence based policy making.”

¹¹⁴ This relates to the official OP proposal submitted to the Commission in May 2007.

initiatives are either underway or planned under the umbrella of a wider public management reform process.

Turning to the issue of the presence of an evaluation society/network, the informal Romanian Evaluation Network *Evalrom* was set-up in July 2006. Evalrom is an open network which brings together professionals (both commissioners and practitioners) in the evaluation of policies, programme and projects from the public and private sectors. Its overall purpose is to “contribute to the inception and consolidation of evaluation culture in Romania”.¹¹⁵ At the time of the case study (end-May 2007), 2 meetings of the network had been held with about the involvement of about 30/35 persons. A working group has been set-up to consider issues around a statute for a possible future formal evaluation association.

As is apparent from the foregoing analysis, the influence of EU evaluation requirements has been a major driver of the development of evaluation capacity in Romania. The new evaluation structures established in the managing authorities are in direct response to the need to meet the evaluation requirements of the Council Regulations.¹¹⁶ In addition, the Phare pre-accession programme has supported the wider PAR process and the various initiatives described earlier to improve the policy making process.

The role of civil society and other players in the evaluation sphere has been very limited in Romania to date, mainly reflecting limited evaluation practice generally.

(7) Conclusions

The reality that Romania acceded to the EU in 2007 and thus has no established track record in the evaluation of SCF programmes places obvious limits on the conclusions that can be drawn as regards evaluation capacity at this early stage. Unlike the other case study countries, Romania can only be assessed in terms of intent rather than actual experience.

That point being made, it seems reasonable to conclude that the Romanian authorities have made a strong commitment to the integration of evaluation practice in both the EU SCF programmes and national policy making domains. In terms of the demand dimension of the benchmarking framework (relating to SCF evaluation), structures in the form of the evaluation units have been put in place and the resources committed (budget and personnel) would appear to be adequate. Coordination and management structures have also been established and the development of on-going or interim evaluation plans is quite advanced. Aspects such as the adequacy and reliability of monitoring systems and data cannot be assessed at this stage.

On the supply side, capacity is clearly quite low (because of low demand in the past) but would appear to be improving as local firms “gear up” to meet expected future evaluation demand. A range of supply-side training and other capacity building initiatives are supported by the Phare programme.

The programmes contain commitments to the dissemination of evaluation outputs and the establishment of mechanisms for handling evaluation outputs and tracking agreed evaluation recommendations.

The wider governance environment has not traditionally been one that was conducive to evaluation practice. The role of evaluation in the policy formulation process is still quite limited in Romania but significant public management reforms are underway which aim to improve the functioning of the policy making process and lend a greater emphasis to *ex ante*

¹¹⁵ See www.evalrom.ro

¹¹⁶ Council Regulation (EC) No 1083/2006 of 11 July 2006.

assessment with integration to the budgetary/resource allocation decisions. Other encouraging developments include the adoption of the National Evaluation Strategy and the role envisaged for the Administrative Capacity OP in supporting these reform initiatives. Outside the public administration, the establishment of *Evalrom* is an encouraging development.

(8) Key Learning Points

With regard to evaluation capacity benchmarking framework

- As is evident from the above analysis, the application of the benchmarking framework did not encounter any serious difficulties in the Romanian context, aside from the obvious fact that evaluation experience is extremely limited. As a result, data or information was simply not available or was very limited on some issues including the quality of monitoring systems, evaluation flow and coverage, evaluation impacts and the role of civil society. However, these are more a product of Romania's "beginner status" as distinct from limitations in the evaluation capacity development benchmarking framework per se.

From Romanian experience

There are a number of interesting and possibly transferable learning points that the analysis of the Romanian situation points to. These include the following:

- The Romanian authorities are using EU funds to good effect in the development of evaluation capacity on both the demand and supply sides (The importance of addressing both demand and supply is emphasised in the literature). This applies to both Phare resources (which are funding some innovative capacity development projects) and the proposed ESF-funded Administrative Capacity OP. This approach may be of interest to both candidate countries and the other new Member States.
- A point highlighted in the literature review is the importance of "champions" –influential actors in key positions who can promote and drive forward the evaluation agenda. This is apparent from the Romanian experience where the Ministry of Economy and Finance (in the SCF sphere) and the GSG (in the national policy domain) respectively have played significant roles in promoting evaluation and enabling wider reforms to the policy making process.

Bibliography

1. Assessment of Evaluation Culture in Romania, Hilary Curley and Eugen Perianu, July 2006
2. National Evaluation Strategy 2007 to 2013, Evaluation Central Unit, Ministry of Public Finance, November 2006
3. Official Proposal for Operational Programme for Administrative Capacity Development, Ministry of Interior and Administrative Reform, May 2007
4. Presentation by Ms. Roxana Mihalache, at UNDP Evaluation Policy Regional Workshop, Istanbul, December 2006
5. Strategy for improving the public policy making, coordination and planning system at the level of the central public administration, General Secretariat to the Government, 2006
6. Evaluation Capacity Development and Institutional Reform in Romania, Wiesner E., UNDP Evaluation Office, 2002