



DG ELARG

EVALUATION GUIDE

November 2008

Directorate E – Evaluation Unit

TABLE OF CONTENTS

1. EVALUATION IN DG ELARG	4
1.1 Evaluation Objectives	4
1.2 Evaluation Mandate	4
1.3 Role of the Evaluation Unit	5
1.4 Monitoring, audit and impact assessment	6
2. PLANNING EVALUATION ACTIVITIES	8
2.1 Evaluation Plan	8
2.2 Planning Individual Evaluations	10
2.3 Ex-ante evaluation	10
2.4 Interim evaluation	11
2.5 Ex-post evaluation	12
2.6 Self-evaluations	12
3. DESIGNING AN EVALUATION	13
3.1 Steering Groups	13
3.2 Terms of Reference	13
3.3 Methodology	14
3.4 Evaluation Questions	16
4. CONDUCTING EVALUATIONS	17
4.1 Four phases of an Evaluation	17
4.2 Format and Presentation of Reports	18
4.3 Quality Assessment of Evaluation Reports	19
5. DISSEMINATION AND USE OF EVALUATION RESULTS	21
5.1 Dissemination and Disclosure of evaluation results	21
5.2 Use and Follow Up of Evaluation Results	23

ANNEX 1:	ELARG EVALUATION MANDATE	25
ANNEX 2:	MISSION STATEMENT OF ELARG EVALUATION UNIT	28
ANNEX 3:	FURTHER DESCRIPTION OF VALUE FOR MONEY CONTROLS	29
ANNEX 4:	PRESENTATION ON DG ELARG EVALUATION FUNCTION	31
ANNEX 5:	EXAMPLE OF TERMS OF REFERENCE	36
ANNEX 6:	EVALUATION CRITERIA	42
ANNEX 7:	MORE ON THE FIVE EVALUATION CRITERIA AND THE ELEMENTS OF INTERVENTION LOGIC	48
ANNEX 8:	RATING FOR QUALITY ASSESSMENT	51
ANNEX 9:	EVALUATION WEBSITE	63
	BIBLIOGRAPHY	65

1. EVALUATION IN DG ELARG

This chapter describes the basic objectives as well as the scope and responsibilities of the evaluation function in DG ELARG.

1.1 Evaluation Objectives

According to the Communication on evaluation¹, evaluation is “*judgement of interventions according to their results, impacts and needs they aim to satisfy*”. The key notion in this definition is that it is a process that *culminates in a judgement (or assessment)* of an intervention. Moreover, the focus of evaluation is first and foremost on *the needs, results and impacts* of an intervention.

The objective of DG ELARG evaluation function is to co-ordinate, manage and implement the evaluation activities of the DG and ensure a proper follow-up of the evaluation findings, results and recommendations. The evaluation function provides inputs for decision-making by contributing to the design of interventions, assisting in an efficient allocation of resources, improving the quality of interventions, and reporting on the accountability of the interventions.

1.2 Evaluation Mandate

DG ELARG Evaluation Mandate sets out the scope and responsibilities of the evaluation function as well as other issues related to independence, objectivity, authority and accountability of the function as well as the evaluation standards (see Annex 1)².

As specified in the mandate the scope of evaluation is to make judgements on interventions according to their results, impacts and needs that they aim to satisfy. To fulfil its responsibilities, the evaluation function shall among others, establish the annual Evaluation Plan, set the guidelines for evaluation and monitoring, provide a tool for quality management and an early warning system for programme implementation, ensure that evaluation results are transmitted effectively to all the relevant stakeholders (country teams, delegations, beneficiary countries) and decision-makers and that those results are taken into account, and provide for an accountability function of pre-accession assistance.

The evaluation function shall ensure that its activities, undertaken either internally or externally, are carried out with sufficient independence and objectivity with respect to operational directorates and programme managers and that the evaluation results are not interfered with. Additionally, the evaluation function is authorised to ask beneficiary countries and members of staff in the Commission to supply in timely manner information deemed necessary for the evaluation activities as well as to report on their use of evaluation results.

¹ SEC(2000) 1051-The Communication on Evaluation.

http://ec.europa.eu/dgs/information_society/evaluation/data/pdf/lib_master/sec2000_1051_strengthening_eval.pdf

² Each Directorate General must have an evaluation function with a clearly defined responsibility for co-ordinating and monitoring evaluation activities of the Directorate General (from the planning of evaluations until their dissemination and use), promoting quality of evaluation and organisational learning, and assisting the central services in the implementation of the Commission Evaluation Policy.

Finally the evaluation function adheres to the evaluation standards set out in the Communication to the Commission "*Responding to Strategic Needs: Reinforcing the use of Evaluation*", of 21 February 2007.

1.3 Role of the Evaluation Unit

According to the Commission Evaluation Standards evaluation activities must be appropriately organised and resourced to meet their purposes. DG ELARG evaluation function is organised as an Evaluation Unit within Directorate E Resources which is clearly visible in the DG's organisation chart and it has clearly defined responsibilities.

To fulfil the mission as set out in the section 1.1, the Evaluation Unit shall³:

- Set guidelines on independent and self-evaluation;
- Establish a multi-annual evaluation programme and an annual evaluation plan consistent with decision-making requirements;
- Provide quality support to decision-making by drawing out lessons learned and feeding them into the design and implementation of programmes, and provide an accountability function of pre-accession assistance;
- Provide value for money controls (prospective and retrospective) in order to ensure the economic, efficient and effective achievement of the objectives of the financial assistance⁴;
- Define quality standards and methodology for evaluation activities;
- Communicate effectively the evaluation results to decision-makers and other stakeholders; regularly monitor the use of evaluation results, and check on how the evaluation recommendations have been taken into account;
- Help develop evaluation capacity in the beneficiary countries;
- Ensure the delivery of and, where appropriate, manage DG ELARG's commitments to evaluation resulting from the Commission's strategic planning and programming cycle and the Communications on evaluation and impact assessment.

The Evaluation Unit is supported in its tasks by the network of evaluation correspondents, appointed by the respective Unit. They contribute to the work of the Evaluation Unit by providing comments to the draft ELARG Evaluation Plan, consultations on evaluation FP and TORs, evaluation guidelines, dissemination of evaluation results, and development of evaluation capacities in beneficiary countries. The network of evaluation correspondents is a useful way of involving operational units closely with the evaluation process.

The Evaluation Unit is supported by the evaluation unit in DG BUDGET, which coordinates the evaluation activities of the Commission and provides methodological and procedural advice to all evaluation functions. For this purpose DG BUDGET coordinates the evaluation network of the evaluation functions of the Commission to

³ See attached Mission Statement in annex 2.

⁴ See attached annex 3, Further Description of Value for Money Controls, section 4.6 of DG ELARG Control Strategy, February 2008.

exchange information and foster best practice. The Secretariat General co-ordinates the impact assessments performed by the DGs. Moreover, DG BUDGET and the Secretariat General have a supporting role to ensure that the evaluation results are fed into the SPP cycle and the budgetary process.

1.4 Monitoring, audit and impact assessment

The activities of the Evaluation Unit need to be distinguished from other complementary but different forms of assessment such as monitoring, audit⁵ and impact assessment.

Monitoring is a continuous and systematic process carried out during the duration of an intervention, which generates quantitative data on the implementation of the intervention, but not usually on its effects. The intention is to correct any deviation from the operational objectives, and thus improve the performance of the programme as well as facilitate subsequent evaluation.

In the specific context of ABM, the Commission has introduced a system of performance monitoring that consists of identifying objectives and indicators for each policy area and activity of a DG and reporting on the attainment of these objectives. This system is designed to provide regular feedback on the implementation of activities (i.e. what outputs have been produced, at what cost (i.e. inputs), over what time period and by whom) and hence a means of assessing the performance of the Commission. Performance monitoring does not however usually collect data about the results and impacts occurring outside the Commission as consequence of its activities. Furthermore, it does not provide answers as to why an activity does or does not attain its objectives, or indeed the relevance of these objectives, neither does it address the question of how performance can be improved. These questions are addressed through exercises of an evaluative nature⁶. Complementarity of the two activities is ensured within the DG since monitoring generates quantitative data on the implementation of the intervention which is fed into the evaluation system as the basis for its judgements of the interventions.

ELARG applies results-oriented monitoring (ROM), which is a form of performance monitoring. ROM is based on regular on-site assessments by monitoring experts of ongoing projects. It focuses on reporting on the inputs, outputs and initial results of project implementation against targets set out in the LogFrames of respective projects. For each project, a short report provides the overall assessment, along with relevant explanations and recommendations for quality improvements. ROM provides important complementary information for internal monitoring by stakeholders and contributes to improving project quality and performance. So far ELARG applies ROM to centralised management.⁷ Complementarity of ROM to evaluations is ensured through its focus on project level monitoring. It is also complementary to evaluation by focusing on outputs produced and the initial results of project performance. The feedback of ROM is fed into the evaluation system.

5 A control function which is primarily concerned with assessing (i) the legality and regularity of project expenditure and income i.e. compliance with laws and regulations and with applicable contractual rules and criteria; (ii) whether project funds have been used efficiently and economically i.e. in accordance with sound financial management; and (iii) whether project funds have been used effectively i.e. for purposes intended.

6 DG Budget: Evaluating EU activities, 2004.

7 Only to the implementation mode of centralised direct management since EAR applied its own type of results-oriented monitoring system. Under DIS the beneficiary countries use also a type of ROM.

Audit, in the public sector, covers a broad range of activities. It comprises of traditional financial audit, which concentrates on inputs and outputs; systems audit, which primarily focuses on the organizations and systems themselves; and performance audit, which may encompass some features of an evaluation. Audit has been extended in scope and within the Commission the internal audit function encompasses the examination and assessment of the adequacy and effectiveness of the DGs' system of internal control.

Both evaluation and audit can be used to assess performance and in this respect some overlap may arise in certain cases. Evaluation and performance audit involve the study of implementation processes and their consequences to provide an assessment of *economy, effectiveness* and *efficiency* of an organisation and/or its activities. However, performance audit tends to be more focused on the implementation of an activity and its immediate effects, while evaluation is centred first and foremost on assessing performance in respect to an intervention's effects. When evaluation examines implementation it normally tries to explain how the results and impacts of an intervention were conditioned by the implementation mechanisms. Furthermore, a broader range of issues fall under the practice of evaluation including an examination of an intervention's *relevance, impact, utility* and *sustainability*⁸.

The evaluation and audit plans within DG ELARG take due account of the complementary but distinctly different roles of evaluation and audit. Evaluation activities focus on judging the relevance, effects and sustainability of the interventions on final beneficiaries. The work of the operational audit unit focuses on examining and assessing the design and performance of systems for managing financial assistance with some emphasis on matters linked to regularity of underlying transactions. Internal audit also concerns itself with design and performance of systems. However, its scope concentrates on the systems of the Commission and Delegations, and it may address all kinds of systems – not only those concerned with financial assistance.

The recently introduced **impact assessment**, which as a rule is now applied to all legislative and policy-defining proposals in the Commission Annual Work programme, has created the need of ensuring maximum coherence and synergy with the evaluation framework in particular with the ex-ante evaluation requirement under the Financial Regulation and its Implementing Rules⁹. This means that two separate exercises are not required. On the other hand the proposals (which are not part of the Commission's Annual Work-programme) for which no impact assessment is carried out, but which have budgetary implications, will continue to require an ex-ante evaluation. It should also be noted that an impact assessment is performed by the operational department concerned.

Although the Commission has traditionally focused evaluation on expenditure programmes, there is an increasing tendency to also evaluate legislation and other non expenditure activities which have substantial impacts on citizens, business and environment. In this respect the new Communication on Evaluation states that the role of evaluation needs to be further developed to help communicate the achievement of policy objectives as well as the challenges in achieving them to decision-makers and stakeholders. Thus evaluation results must be used not only for improving preparation, implementation and performance of individual policy instruments but also as an input for setting political priorities. Complementarity of impact assessment and evaluation of non-

8 DG Budget: Evaluating EU activities, 2004.

9 All proposals for programmes or activities occasioning expenditure or a reduction in revenue for the budget shall be the subject of an ex-ante evaluation.

expenditure programmes is ensured since ex-post evaluations are used as inputs for impact assessment.

2. PLANNING EVALUATION ACTIVITIES

This chapter starts with a description of how the Evaluation Plan is prepared including decisions on its coverage, and the ways it is reported on. Then it deals with different types of evaluations carried out within DG ELARG and in what way they are contracted. It further explains DG ELARG's approach of results-oriented evaluation.

2.1 Evaluation Plan

In accordance with the Commission Evaluation Standards, the Evaluation Unit prepares an annual evaluation plan and an indicative multi-annual evaluation programme in consultation with the other units in the Directorate General. The Evaluation Plan ensures that evaluations are chosen and undertaken in a transparent and timely manner so that they provide the decision makers with timely and relevant information which contributes to the design of the interventions and assists in an efficient allocation of resources. The Evaluation Plan is attached to the Annual Management Plan document.

The Evaluation Plan includes the specific evaluation projects as well as other horizontal activities carried out by the DG ELARG Evaluation Unit. It also includes the Self Evaluation Programme (at project level) carried out by the operational departments/delegations concerned.¹⁰ It comprises an estimate of the human and financial resources needed to bring it into effect.

The Evaluation Plan is prepared once a year by the Evaluation Unit in consultation with the other units in DG ELARG, the Delegations and other relevant services. It is submitted for approval to the Director General. Once approved it becomes an annex to DG ELARG Annual Management Plan. Thereafter the evaluation function is responsible for co-ordinating and monitoring the implementation of the Evaluation Plan and reporting on its achievements to the Director General.

DG ELARG Evaluation Plan takes due account of the legal obligations of the Financial Regulation and Implementing Rules that require interim and/or ex-post evaluation of all programmes and activities exceeding €5 million. In addition, all proposals for expenditure programmes shall be subject to an ex-ante evaluation. The evaluation results shall be disseminated to the spending, legislative and budgetary authorities. Moreover, the Commission binding evaluation standards require that all activities (spending programmes and non-spending activities) addressed to external parties must be periodically evaluated in proportions to allocated resources and expected impact. The IPA implementing rules set out that all programmes shall be subject to interim and/or ex-post evaluation. Finally, the JMC mandate (for decentralised management) requires preparation of sector interim evaluations and a country summary evaluation report before each JMC meeting.

¹⁰ See section 2.6 page 11.

The Evaluation Plan is drawn up to support DG ELARG results framework of financial assistance.¹¹ DG ELARG results-oriented evaluation¹² is mainly concerned with reporting on the results and impacts achieved by operations, and distils the operational experience into knowledge of what works and why against DG ELARG results framework of financial assistance.¹³

The Evaluation Plan is also closely related to the results framework of the Strategic Planning and Programming Cycle (SPP cycle), which aims to link decisions on political priorities and related resources. The evaluation planning and the evaluation findings figure strongly at the four stages of the SPP cycle - Annual Policy Strategy (APS) decision, preliminary Draft Budget (PDB), Annual Management Plan (AMP) and Annual Activity Report (AAR). As an example, it is drawn up to ensure the relevant coverage of the DG activities taking into account the key risk factors of the pre-accession assistance, as specified in the AAR.

The Plan also addresses the value-for-money related risks identified in previous evaluations, such as the development of a results-based framework of financial assistance¹⁴, weaknesses in strategy formulation and objectives, lack of programming readiness and risks of implementation related to low performing sectors.

Reporting on the Evaluation Plan occurs three times a year in the context of AOSD reporting which is related to the Annual Management Plan objectives and identification of any risk which may affect them. Additionally at the end of the year in the context of the preparation of the Annual Activity Reports, the Evaluation Unit must report on the degree of achievement of the objectives fixed in the annual Evaluation Plan.

Finally, the Annual Evaluation Review (AER) compiled by DG BUDG provides a comprehensive overview of the evaluation reports across the Commission services highlighting key evaluation findings on the effectiveness, efficiency or relevance of EU activities. The first part of the AER containing evaluation results and lessons learned is forwarded by DG BUDG to the European Parliament and the Council within the framework of the Preliminary Draft Budget. Further use of this document is made by the Committee of Budgetary Control of the European Parliament as well as by the CoA.

11 For a presentation on the evaluation function in DG ELARG, see annex 4.

12 DG ELARG results framework of evaluation involves the features of independent evaluations, participatory approach, resident evaluators (interim evaluation scheme), results-oriented monitoring, quick turnaround products, systematic follow –up (e-manager, monitoring committees), greater use of prospective and ongoing evaluations, focus on sector, thematic and country evaluations (independent), self-evaluations (project level), quality of evaluations (quality assessment grids and quality assurance groups) and local evaluation capacity development.

13 It ensures that evaluations are chosen to comply with DG ELARG's approach of results-oriented evaluation.

14 ELARG results-based frameworks of financial assistance consist of targets and results-based indicators set out in Annual Progress Reports, NPAAs, MIPDs, Annual Programmes and LogFrames of PFs. Progress on achieving these targets are reported on in monitoring reports, AOSD/SPP Cycle, QSG and JMC reviews. The recent introduction of this results-based framework has enabled ELARG to reinforce its approach of results-oriented evaluation. The Evaluation Unit provides for regular support to preparing results-based indicators in the MIPDs as well as to the preparation of LogFrames.

2.2 Planning Individual Evaluations

The evaluation can be categorised according to different criteria, such as temporal scope (ex-ante, interim, ex-post)¹⁵, how it is conducted (internal/external)¹⁶ or its purpose (summative/formative)¹⁷. Each evaluation should be carefully managed for the purpose of its own efficiency.

DG ELARG evaluation activities can be divided into two main groups: independent evaluations and self-evaluations. Within the independent evaluations there are three different types: ex-ante, interim and ex post. As part of DG ELARG's results-focused approach to evaluation, it makes greater use of ex-ante and interim evaluations.

In accordance with PRAG guidelines evaluation contracts are designed "to allow a maximum contract size in order to achieve economies of scale, maximum co-ordination and simplify administration". Larger contracts above € 200.000 are procured through open/restricted tender procedure. Where appropriate, for contracts below € 200.000, as for ad hoc evaluations, the framework contract procedure is used.

2.3 Ex-ante evaluation

An ex-ante evaluation can take place at different levels of activity addressing a policy, programme or project. They serve mainly as an aid to programming by improving the quality, relevance and comprehensiveness of programme design, and are usually carried out in parallel to the programme preparation. An ex-ante evaluation should be an analytical and iterative learning process not dissociated from the process of preparation and design of an intervention.

DG ELARG ex ante evaluations are normally carried out in house by the Evaluation Unit involving operational staff. It is important to carry out the evaluation in tandem with the operational units to facilitate direct feedback of the evaluation results into the design of the intervention/preparation of the proposal. However, even if the co-ordination and main part of the ex-ante analysis is carried out in-house, some parts of it can often be entrusted to an external expert. In this case this expertise is normally recruited through framework contracts.

The ex-ante evaluations should start early in the process when options for improvement are still open. As some elements of the proposal may change in the course of its development, it is often necessary to revise some parts of the analysis accordingly.

As an example, the ex-ante evaluations of the MIPDs provided concrete recommendation to the operational units in charge of the preparation of these documents. The findings of the evaluation helped to reinforce the treatment of programming readiness in the MIPDs, and made a substantial contribution to the deliberations and the results of the Quality Support Group process and subsequent

15 Temporal scope of evaluations can also be categorized as prospective (ex-ante) and retrospective (interim and ex-post) evaluations.

16 Different evaluation situations are more or less suited to internal or external evaluation. The use of contractors is particularly suitable for summative evaluations, since external evaluators may be more independent in their assessment. On the other hand, internal evaluations carried out by the evaluation function and involving staff is particularly useful where the formative dimension is dominant.

17 Summative evaluation is mainly concerned with determining the effectiveness of an intervention for the purposes of accountability. Formative evaluation is concerned with examining ways of improving and enhancing the implementation and management of interventions. Evaluations often combine both formative and summative dimensions.

revision of the MIPD guidelines/documents. Moreover, the written QSG assessments of the annual programmes presented at the QSG serve as a form of ex-ante evaluation.

More information on how to carry out an ex ante-evaluation can be found in DG BUDGET's *Ex- ante Evaluation Guide*¹⁸.

2.4 Interim evaluation

Interim evaluation (also referred to as ongoing, intermediate and mid-term evaluation) examines an ongoing activity whether this is a programme of limited duration or a policy, which continues for an indefinite period. An interim evaluation has an important role to play in producing direct feedback into the implementation process and thus help to improve the quality of ongoing interventions. Moreover, since new initiatives are often prepared long in advance, interim evaluation is also a very important source of information for the design purposes for the next generation of a programme, new policy orientations, etc.

DG ELARG Evaluation Unit undertakes ad hoc interim evaluations in the countries where pre-accession assistance is managed centrally by the Commission headquarters or where it has been deconcentrated (devolved) to the Delegations. The European Agency for Reconstruction (EAR) managed pre-accession assistance in Serbia (including Kosovo), Montenegro and the Former Yugoslav Republic of Macedonia in a centralised indirect implementation mode, and was therefore also responsible for carrying out the interim evaluations in those countries. In December 2008 the Agency was closed down and its responsibilities were transferred over to the EC Delegations (or in the case of Kosovo, the EC Liaison Office).

DG ELARG interim ad hoc evaluations are usually carried out by external contractors operating under the responsibility of the Evaluation Unit through framework contracts.

In the countries where the management of pre-accession assistance has been decentralised (decentralised management with ex-ante controls, DIS), DG ELARG has set up a centralised Interim Evaluation scheme, which essentially constitutes a programme management tool to provide programme managers (Country Teams, Delegations and co-ordinators in the candidate countries) with assessments of the state of implementation of their programmes including programme performance, efficiency and sustainability against stated objectives, and lessons learned with a view to improving programme implementation as well as the design of future programmes. In contrast with the ad hoc evaluations, the Interim Evaluation scheme also contains provisions for building up evaluation capacity in the candidate countries.

The interim evaluation scheme is externally contracted, awarded through a tender evaluation in order to achieve economies of scale, maximum co-ordination and simplify administration. Under this scheme a team of external evaluators is resident in the candidate countries and in Brussels under the supervision of the Evaluation Unit. For in-depth interim evaluations or evaluations of more complex/technical nature, the team of external evaluators is supported by short-term experts.

More information on the Interim Evaluation Scheme can be found in DG ELARG "*Phare IE guide*"¹⁹. Under decentralised management without ex-ante controls, the responsibility for conducting interim evaluations is decentralised to the National Authorities. The evaluation principles as set out in the Interim Evaluation Guide shall be

18 http://ec.europa.eu/budget/library/documents/evaluation/guides/ex_ante_guide_2001_en.pdf

19 http://iewebsite/docs/ie_guide/IEguide%20_Part_I_revised_07_12_04.pdf

respected by the National Authorities but the practical evaluation arrangements can be adjusted to suit the particular needs of the beneficiary country.

2.5 Ex-post evaluation

Ex-post evaluation is conducted either on or after completion of an intervention. It embraces the entire intervention period, with a focus on the impacts, efficiency and effectiveness of the intervention. It should also assess how sustainable the realised impacts are and what are the main factors behind success or failure of an intervention. Since it usually takes time for impacts to materialise, ex-post evaluation should ideally be carried out some time after the expiry of an intervention.

DG ELARG ex-post evaluations review the performance of pre-accession assistance in beneficiary countries. This helps to draw conclusions and make recommendations that should influence the shape and magnitude of pre-accession assistance for the remaining and future candidate countries. In this respect DG ELARG ex-post evaluations provide accountability as to the value for money and the use of community funds, and provide lessons learned for decision-making on improvements of pre-accession aid for the future candidate and potential candidate countries.

Ex-post evaluations are performed by the Evaluation Unit. They are normally undertaken through an external contractor. This helps to strengthen the credibility of the evaluation exercise from an accountability perspective by underpinning the independence, objectivity and transparency of the process. It also helps to optimise the use of scarce human resources and brings in expertise, which might not be available within the Commission. DG ELARG ex-post evaluations tend to be large exercises and they are usually contracted through a tender in order to allow a maximum contract size and to achieve economies of scale, maximum co-ordination and simplified administration.

2.6 Self-evaluations

Self-evaluations are complementary to independent evaluations. They are limited to project level and conducted under the responsibility of the operational departments during the preparation, implementation and/or completion of an intervention. Self-evaluation is a systematic and participatory exercise for the stakeholders of a project aimed at judging interventions and ensuring the lessons learned are fed into the whole cycle of the project. They can also be used as input for the independent evaluations.

Since self-evaluation is conducted mainly for the benefit of the project managers, it should be commissioned by them. Self-evaluation of centralised and deconcentrated projects should be commissioned by Commission headquarters and EC delegations respectively. For decentralised projects, various forms of partnership between headquarters, EC delegations and beneficiaries can be envisaged.

DG ELARG Evaluation Unit performs an advisory function during the self-evaluation process, including consultations on work programmes and terms of reference. It also provides technical and methodological support in the development and application of self-evaluation tools and carries out ex-post control checks on samples of self-evaluations.

Self-evaluations can be conducted fully as in-house evaluation, or supported by a contractor, or be fully externally contracted.

The Evaluation Unit of DG ELARG has issued, in September 2006, a "*Practical Guide for DG Enlargement Self-evaluations*"²⁰ which sets out the guidelines for designing and conducting a self-evaluation and also deal with the reporting, dissemination, use and follow-up of self-evaluations results.

3. DESIGNING AN EVALUATION

This chapter deals with the design of an evaluation, thus how the steering group is set up and how the key specifications for the work to be carried out are defined and specified in the Terms of Reference.

3.1 Steering Groups

The Steering Group is usually composed of a group of officials, who are directly concerned and/or involved with the intervention subject to the evaluation and furthermore have some knowledge of evaluation.

The principal role of the Steering Group is to ensure a high quality and useful evaluation and entails different activities such as participating in the development of the terms of reference, taking part in the assessment of the quality of the evaluation work at key moments, facilitating access to information, giving an opinion on the quality of the final report to the Evaluation Unit, and assisting in the feedback of the findings and recommendations of the evaluation exercises. The members of the group should not attempt to influence the evaluator to omit certain evidence or to come to conclusions that are not substantiated by the evaluation evidence.

A Steering Group can make it easier to get access to information and increase understanding of the facts and events of an intervention and contribute to a quicker dissemination of conclusions and recommendations. It can also help to a better acceptance of the evaluation and its findings by the stakeholders.

Appropriate Steering Group arrangements are decided by DG ELARG Evaluation Unit for each evaluation exercise. Usually the Steering Group is composed of a group of officials mainly from DG ELARG and beneficiary countries, under the chairmanship of the Evaluation Unit.

The process for the Interim Evaluation is participatory by systematically involving beneficiaries and stakeholders from planning the evaluation to implementation and reporting. The evaluations in the new Member States are carried out in partnership with the evaluation functions in the respective countries that become member of the steering group for such an exercise.

3.2 Terms of Reference

The Terms of Reference (TOR) is the key document in the evaluation process. It provides all the key specifications for the work to be carried out including the basis on which the evaluation method is designed.

The main elements which shall be included in every TOR are the following:

²⁰ http://iewebsite/docs/ie_guide/Self-Evaluation%20Guide%20-%20FINAL%2012Sep06.doc

- ▶ *Objectives and scope.* This part of the TOR explains the purpose and the scope of the evaluation, and thus defines the general objectives and intended use of the evaluation project.
- ▶ *Background and context.* This section should set out an identification of the key issues that serve as a relevant background and context for the evaluation.
- ▶ *Evaluation questions.* Against the identification of issues in the background and context section above, a set of evaluation questions should be drawn up aimed at directing the work of the evaluator. For further details see section 3.3.
- ▶ *Methodology.* The approach for data collection and analysis can contain general orientations on methods, typologies to be applied as well as minimum requirements concerning data collection. This methodological part fulfils four major tasks, which are implemented largely in a sequential manner: 1) providing a focus and structure to the evaluation project; 2) collecting data; 3) analysing data collected and from other sources; and 4) providing judgements.
- ▶ *Reporting and target audiences.* The format of reporting and the target audiences should be specified in the TOR.
- ▶ *Activities, resources and timetable.* Resource requirements and timetable for the evaluation exercise should be specified in the TOR.

The draft TOR for independent evaluations is prepared by the Evaluation Unit. In the case of an external evaluation, the TOR must also include the contractual, financial and administrative clauses and quality criteria. The draft TOR is consulted with the relevant stakeholders and beneficiaries. The Steering Group shall participate in the validation of the TOR that is then finalised and issued by the Evaluation Unit.

As to self-evaluations, the project manager in charge of carrying out the self-evaluation exercise prepares the TOR. However he/she shall consult DG ELARG Evaluation Unit to ensure that the main elements specified above are included.

For further details on the preparation of a TOR, see also “*Evaluating EU Activities. A Practical Guide for the Commission Services*” by DG BUDGET. An example of a TOR for an evaluation is in Annex 5.

3.3 Methodology

3.3.1 Evaluation criteria

The basic concepts and the evaluation methodology applied to the various types of evaluation are based on five key evaluation criteria: relevance, efficiency, effectiveness, impact and sustainability. The methodology is consistent with common practices of evaluation and reflects the elements and concepts of intervention logic that analyses a project along its inputs, outputs and effects.

▶ Relevance

The relevance of a project relates primarily to its design and concerns (a) the extent to which its stated objectives correctly address the identified problems or real needs and (b) the extent to which the project is feasible, taking account of the resources and capacities of the beneficiary. It has to be kept under review throughout the life of the project in case changes occur either in the nature of the very problems originally identified, or in the circumstances - whether physical, political, economic, social,

environmental, institutional or policy - in which the project takes place, necessitating a corresponding change of focus.

► Efficiency

The efficiency criterion deals with the relationship between resources employed and results achieved. A key question it asks is “have things been done right?” and thereby also addresses value-for-money, that is whether similar results could have been achieved better by other means at lower cost within the same amount of time.

► Effectiveness

The effectiveness criterion concerns the extent to which the project is meeting its operational objectives by delivering the required activities and outputs and achieving the intended results. A key question is “what is in place that was not there before?” e.g. equipment installed, teachers trained, procedural manuals”. Another key question is what difference the project made in practice, as measured by how far the intended direct beneficiaries really benefited from the products or services it made available.

► Impact

The impact, sometimes referred to as outcome, is threefold.

The *immediate impact* concerns results, i.e. the extent to which assistance has delivered substantive, sustainable change to the direct beneficiaries of the assistance. What has changed in the way that direct beneficiary institutions and systems function as a result of the assistance (i.e. as a result of the activities and outputs identified under “effectiveness”)?

The *intermediate impact* relates to the extent to which the project has influenced, or is likely to influence, people’s behaviour beyond the envisaged scope of the direct beneficiaries of the project, including through unplanned project “spin-offs”: e.g. institutions other than the direct beneficiaries requesting similar assistance, additional related activities planned and/or undertaken by the project beneficiary outside the scope of the project.

The *wider impact* concerns the extent to which immediate and intermediate impacts are contributing to the wider objectives of the project. What else needs to be done to achieve the wider objectives, and is it being done, or likely to be done in the foreseeable future? To what extent have the benefits received by the target beneficiaries had a wider overall effect on larger numbers of people in the sector or region or on society as a whole?

► Sustainability

The final criterion, sustainability, relates to whether the positive impacts of the project, whether immediate, intermediate or wider, are likely to continue after external funding ends, and be sustained at the level of the sector, region or society as a whole.

Annexes 6 and 7 elaborate further on the five evaluation criteria in relation with the intervention logic.

3.3.2 Performance rating

Performance ratings are in accordance with the following 6 rating categories:

Highly satisfactory (HS):	The programme/ project achieved/is likely to achieve at least acceptable progress toward all major relevant objectives. No major shortcomings were identified.
Satisfactory (S):	The programme/project achieved/is likely to achieve acceptable progress toward all major relevant objectives. No major shortcomings were identified.
Moderately Satisfactory (MS):	The programme/project achieved/is likely to achieve acceptable progress toward <i>most</i> of its major relevant objectives. No major shortcomings were identified.
Moderately Unsatisfactory (MU):	The programme/project did not/is not likely to make acceptable progress on <i>most</i> of its major relevant objectives.
Unsatisfactory (U):	The programme/project did not/is not likely to make acceptable progress toward <i>most</i> of its major relevant objectives.
Highly unsatisfactory (HU):	The programme/project did not/is not likely to make acceptable progress toward <i>any</i> of its relevant objectives.

In the situation where the evaluator considers that there is insufficient evidence to enable a rating to be made for an evaluation criterion, the box should be marked 'NR' i.e. 'No Rating Possible'²¹

3.4 Evaluation Questions

The elaboration of evaluation questions is a core element of the TOR and essentially aims at directing the work of the evaluator by focusing the evaluation work on a limited number of key points.

²¹ For additional information please see the Phare Interim Evaluation Guide.

The issues that an evaluation should address depend to some extent on the type of evaluation whether it is ex-ante, ex-post or interim. In an ex-ante and also interim evaluation, relevance, efficiency, effectiveness, *likely* impact and *likely* sustainability are the most relevant evaluation issues and therefore the evaluation questions shall be addressed in this respect. Ex-post evaluations will rather focus on questions of impact and sustainability. The choice of the evaluation questions shall be made taking into account its usefulness for the evaluation as well as its feasibility.

In the area of financial assistance the evaluation questions frequently address performance and thematic/cross-cutting issues. As to performance evaluation questions, one of the key issues relates to the performance of the intervention, and is addressed by performance evaluation questions that reflect the five key evaluation criteria of relevance, efficiency, effectiveness, impact and sustainability. Such standard evaluation questions deal with: needs assessment and design of the intervention; how economically inputs/ resources have produced outputs; the extent to which outputs were delivered according to operational objectives; the extent to which immediate, intermediate, wider impacts were, or are likely to be, delivered according to respectively immediate, intermediate and overall objectives; and the long-term viability of the results/impacts following project completion. In addition to these performance evaluation questions, the evaluation may also include questions related to thematic /cross-cutting issues.

The inputs from all the stakeholders concerned shall be taken into account when drafting the evaluation questions. Therefore evaluation questions are established by the Evaluation Unit after having consulted the different stakeholders, for example by inviting them to a kick-off meeting where the draft evaluation questions are discussed.

4. CONDUCTING EVALUATIONS

This chapter addresses the different arrangements related to conducting an evaluation, discussing firstly the four phases of an evaluation. Secondly it addresses the issues related to format and presentation of reports followed by quality assessment of evaluation reports.

4.1 Four phases of an Evaluation

When the evaluation is undertaken internally, the Evaluation Unit (or the project manager, for self-evaluations) is responsible of conducting the evaluation.

As to external evaluation when the procurement process has been finalised and the evaluator has been contracted, the implementing phase will be managed by the contractor. The contractor nominates a project director. The external evaluation exercise is supervised by the Evaluation Unit.

The methodology of an evaluation is usually composed of a combination of tools and techniques assembled and implemented in order to provide answers to the questions posed within the framework of an evaluation, with due regard to the time, budgetary and data constraints of the project.

An evaluation, in particular interim and ex-post can be broken down in four phases of execution:

► Structuring phase

The structuring phase embraces elaboration on the evaluation questions, indicators of achievement, evaluation methodology and the organisation of the evaluation activities, including kick-off meetings, seminars, debriefing meetings and final presentation and culminates with the Inception Report. The TOR is prepared by the Evaluation Unit (a project manager for a self-evaluation) and the inception report prepared by the contractor is approved by the Evaluation Unit.

► Data collection phase

The data collection phase in the field commences after the inception report has been delivered by the contractor and approved. It consists firstly of deciding from which actors to collect data on implementation and effects and what questions to ask them, and then of actually implementing the data collecting strategy, using interviews, and questionnaire surveys, etc. Data collected will usually be of both a qualitative and quantitative nature, and will provide factual information along with perceptions and opinions. The contractor has already in their technical proposal suggested an appropriate methodological approach accepted by the Evaluation Unit.

► Analysis of data

The analysis of data firstly takes place at the level of each data collection tool, for example a statistical analysis of survey data, coding and abstraction of data from interviews. The cross-analysis of data from different data sources including secondary data constitutes the second step of this phase. An essential part of work at this point is an analysis of causality and the attribution of effects to the intervention.

► Formulation of judgements

Formulation of judgements is the final stage of the evaluation. Once all the data have been analysed and the evaluation results compiled the success of an intervention can be judged. It involves in particular assessing an intervention in respect to specific evaluation questions that reflect generic evaluation issues such as relevance, effectiveness, efficiency, impact, sustainability, as well as well defined judgement criteria.²²

The tools and techniques that will be used during the evaluation also depend on whether the questions to be addressed are forward looking (ex-ante evaluations) or retrospective (interim and ex-post) in nature since some of the tools and techniques may be better suited to one or the other of these situations. The Evaluation Unit is involved during all the four phases.

4.2 Format and Presentation of Reports

The following text on format and presentation relates mainly to interim, ex-post and self-evaluations. Ex-ante evaluation reports shall respect the requirements of the article 28 of the Financial Regulation²³ as well as DG BUDGET guidelines for ex-ante evaluation.

²² A judgement criterion specifies an aspect of the evaluated intervention that will allow its merits or success to be assessed. It is used to answer an evaluation question positively or negatively. Whilst "judgement criterion" is the appropriate word, an acceptable alternative is "reasoned assessment criterion."

²³ Any proposal submitted to the legislative authority which have an impact on the budget, including changes in the number of posts, must be accompanied by a financial statement and the evaluation provided for in Article 27 (4).

Evaluation reports will be most useful to stakeholders and others if they are substantive, concise and above all, readable. In general, the main sections of the Report are as follows:

Executive Summary: the report must have an executive summary presented in the form of a synthesis. It should be short, no more than five pages, a tightly drafted, to the point stand-alone document. It must bring out all the key issues/analytical points and key messages/specific recommendations. The key recommendations must carefully be targeted to the appropriate audiences at all levels.

Main text: starts with an introduction describing the background and scope of evaluation. The body or core of the report provides for quantitative and qualitative information analysed according to the five evaluation criteria, describing the facts and interpreting or analysing them in accordance with the key questions pertinent to each criterion. The core of the report may also include an analysis of thematic/ cross-cutting issues.

Conclusions: they stem logically from findings and are based on impartial judgement criteria. Conclusions are properly addressed to the evaluation questions and coherently and logically substantiated by evaluation findings. They are orderly presented and related (categorised, ranked, priorities, sequence). Conclusions are free of personal or partisan considerations.

Recommendations: they stem logically from conclusions and cover all relevant main conclusions. The ultimate value of an evaluation depends on the quality and credibility of the recommendations offered. Recommendations should therefore be as realistic and pragmatic as possible.

Recommendations should be presented in tabular format reflecting the following parameters: recommendation/action, name/title of addressee and deadline for implementation. Recommendations from previous if any evaluation reports that are still valid should be added separately.

4.3 Quality Assessment of Evaluation Reports

A system for quality assessment is needed in order to ensure that all evaluation reports come up to quality standards in the field as part of DG ELARG results framework of evaluation. A high quality demand is also important for establishing the future quality standards within the emerging evaluation communities in the candidate countries. It is important to stress that it is the contractor's full responsibility to ensure the quality of all the deliverables included in the contract.

4.3.1 Validation of Evaluation Reports

During the evaluation process the external evaluator will deliver a number of various reports at various stages of the evaluation.

Inception Report

The Inception Report draws the structuring phase of an evaluation. It records the activities carried out during the inception period and sets out a detailed planning of the project's activities, including the work programme for the entire project. It also describes in detail how the method proposed by the evaluator is going to be implemented in the light of examination of the quality and appropriateness of existing data, and in particular

how the method will answer evaluation questions and provide a judgement. It provides the Evaluation Unit and the Steering Group with the opportunity to check the feasibility of the method proposed and the extent to which it corresponds with the information needs outlined in the Terms of Reference

First Draft report

The Draft report will provide the conclusions of the evaluator in respect to the evaluation questions in the Terms of Reference. The conclusions will be clearly based on evidence generated through the evaluation. Judgement provided should be clear and explicit.

For interim, ex-post and ex-ante evaluation, the quality of the documents is reviewed by the evaluation function, and when applicable also the Steering Groups.

For the interim evaluation the first draft of report is issued for comments simultaneously to all parties involved: ECD, NAC, NIPAC, DG ELARG's country teams, implementing bodies, and other national authorities. The Evaluation Unit verifies the report's overall compliance with the guidelines, template and procedures. Written comments must reach the evaluator within 10 working days of the issue of the draft. Comments should be specific and enable the evaluator to modify the text directly.

For the ex-post evaluation the first draft is issued for comments to the Evaluation Unit. The task manager in the Evaluation Unit makes written, specific comments and recommendations to the evaluator. The evaluator has to judge the comments and recommendations and make modifications. The modifications will be presented in a table for treatment of recommendations including actions taken. If no actions have been taken there should be an explanation of the reasons.

The contractor shall arrange debriefing meetings in Brussels and/or in benefiting countries on the draft reports in co-ordination with the Evaluation Unit.

For the self-evaluation, the project manager or the external contractor produces a draft, on the basis of its analysis. Thereafter the draft is issued for comments simultaneously to all stakeholders involved. The draft is forwarded to the Evaluation Unit for information.

Final report

The final evaluation report takes into account the results of quality assessments and discussions with the Steering Group insofar as they do not interfere with the independence of the evaluator in respect to his/her opinions.

As to interim evaluation, the Evaluation Unit receives a revised report together with a table indicating how the various comments received have been handled. The evaluator makes his own judgement as to whether to take comments or modifications into account, or to simply include them as dissenting views in an annex to the report. The Evaluation Unit finally checks the report against the formal requirements of the Guide, and checks in particular whether the comments have been handled properly. The Evaluation Unit then updates its quality rating.

As to ex-post evaluation, the Evaluation Unit receives a revised report together with a table indicating how the comments have been handled. The Evaluation Unit checks the report and in particular whether the comments have been handled properly. Before the contractor produces a final report the Evaluation Unit forwards the draft for consultation to stakeholders within the Commission such as the country teams and co-ordination directorates of DG ELARG, Delegations, and relevant line DGs. Further stakeholders for consultations are the NACs/NIPACs and implementing agencies in the beneficiary countries.

As to self-evaluation, the project manager concerned for an internal evaluation makes his/her own judgement as to whether to take into account the comments received on the draft report, or to simply include them as dissenting views in an annex. In the case of an external self-evaluation, the revised report is sent by the contractor to the project manager together with a table indicating how the various comments received have been handled. The project manager checks the report against the formal requirements of the Guide, and checks in particular whether the comments have been handled properly.

The quality of all the reports shall be in accordance with the Commission's binding Evaluation Standards²⁴.

4.3.2 Quality Assurance Group

The contractor of an external evaluation establishes a suitable quality control system of their outputs that in turn is subject to the quality assessment grid established by the Evaluation Unit.²⁵ As part of DG ELARG results framework of evaluations, a Quality Assurance Group is established to assist in the quality assurance process, more particularly in terms of quality control checks of a sample of the interim and ex-post evaluation reports and of the methodology applied.

The Quality Assurance Group comprises of 2-3 independent evaluation experts and meets on request by the Evaluation Unit. The Group helps to ensure that the state-of-art of the evaluation methodology is adequately applied and that both the evaluation process and the reports are of high quality.

5. DISSEMINATION AND USE OF EVALUATION RESULTS

This chapter addresses the feedback mechanisms related to evaluation findings at the end of the evaluation project. These mechanisms relate to final reporting, dissemination and use of evaluation results. The final reporting takes place when the contractor transmits the evaluation report to the Evaluation Unit/Steering Group. Dissemination, on the other hand, refers to the set of activities by which evaluation findings are made available to the target group as well as to the general public.

5.1 Dissemination and Disclosure of evaluation results

Evaluation results are communicated in such a way to support management and decision-making processes to ensure the maximum use of the results and that they meet the needs of decision-makers and stakeholders.²⁶

Ex-ante, interim and ex-post evaluations reporting is targeted to a wide spectrum of audiences at all levels:

- Task managers in the NAC/NIPAC or Implementing Agencies and Delegations
- DG ELARG Country Teams
- JMCs and SMSCs under Phare; under IPA and for candidate countries there is an IPA Monitoring Committee and Sectoral Monitoring Committee per component

24 Communication to the Commission from Ms Grybauskaite in agreement with the President, SEC(2007)213.

25 See attached annex 8 Rating for Quality Assessment, for a guide on scoring the criteria and the quality assessment form.

26 Communication to the Commission from Ms Grybauskaite in agreement with the President, SEC(2007)213.

- Senior management (authorities in the candidate country, Commission services)
- Court of Auditors
- National audit offices(or equivalent) of the candidate countries
- Member States delegations to the IPA Management Committee

Self-evaluation is targeted to a more narrow spectrum of audiences:

- Project managers or Implementing Agencies and Delegations
- DG ELARG Country Team
- DG ELARG Evaluation Unit

DG ELARG has in place different mechanisms to disseminate the evaluation results:

- ▶ Ex-ante evaluation results are firstly debriefed, discussed and followed up through direct contacts with the operational units in charge of the evaluation. Additionally the results of the ex-ante evaluation are also discussed at the QSG.
- ▶ Regular interim and self-evaluations are presented, once a report is issued, in a debriefing meeting convened by the external contractor/project manager concerned and attended by all stakeholders. The main purpose is to jointly assess the achievements of the project evaluated (production of outputs, achievement of objective and outcomes) and to discuss the report's recommendations. The meeting focuses on the means of and the timing for implementing the recommendations.
- ▶ Within a month of the debriefing meeting, the follow-up table signed by the chairperson of the debriefing meeting is sent back to the external contractor /project manager concerned. The follow-up table describes the actions to be taken by each stakeholder to implement the recommendations. It also describes possible dissenting views of the stakeholders.²⁷
- ▶ Thematic, consolidated and ex-post evaluation results are communicated via meetings and evaluation seminars, debriefed at evaluation seminars, in which stakeholders are invited to participate and to comment on the evaluation findings.
- ▶ Finally, the Annual Evaluation Review (AER) compiled by DG BUDG provides a comprehensive overview of the evaluation reports across the Commission services highlighting key evaluation findings on the effectiveness, efficiency or relevance of EU activities. The first part of the AER containing evaluation results and lessons learned is forwarded by DG BUDG to the European Parliament and the Council within the framework of the Preliminary Draft Budget. Further use of this document is made by the Committee of Budgetary Control of the European Parliament as well as by the CoA.

According to the Commission's binding evaluation standards²⁸, the evaluation results must be publicly available and targeted summary information should be prepared to

²⁷ DG ELARG, Self –Evaluation, p.10.

²⁸ Communication to the Commission from Ms Grybauskaite in agreement with the President, SEC(2007)213.

facilitate communication to the general public²⁹. Moreover, according to the Financial Regulation (art.27.4) ex-ante and ex-post evaluation results shall be disseminated to spending, legislative and budgetary authorities³⁰.

The disclosure arrangements at DG ELARG are organised as follows:

- ▶ Final ex-post evaluation reports are disclosed to the general public.³¹ The publication of these reports is authorized through a cover note by the Head of Unit or the Director for Resources.
- ▶ Thematic Interim Evaluation reports that constitute a final interim evaluation are disclosed to the general public. The publication of these reports is authorized through a cover note by the Head of Unit.
- ▶ Interim evaluations (sector and country evaluations) are disclosed to all the internal stakeholders including the beneficiaries. Executive summaries of these reports are also disclosed to Member State representatives in the IPA Management Committee. But, since most of the sector interim evaluations (and thus country interim evaluations) normally represent only preliminary assessments of project implementation of financial assistance frequently at very early stages, it would be premature to disclose the evaluation results of these reports to the general public. A consolidated report of sector and country interim evaluations is periodically disclosed to the general public. The publication of these reports is authorized through a cover note by the Director for Resources.
- ▶ Evaluation results of final ex-ante evaluations of programmes and activities which entail significant spending are disseminated to spending, legislative and budgetary authorities.
- ▶ Since self-evaluation is conducted at non-programme level (and thus not bound by the evaluation standards) and essentially serves as a project management tool aiming at supporting the internal stakeholders, it does not require that such reports are disclosed to the general public.

Country, sectoral and thematic reports are published on the DG ELARG Evaluation website (see annex 9). The evaluation reports which are published on the ELARG intranet will also be available on the T:drive. ELARG will launch a feasibility study to introduce interactive country maps to obtain information including evaluation results. Moreover, a register of lessons learned would be set up including a search function on sector, thematic and country issues.

5.2 Use and Follow Up of Evaluation Results

The Evaluation Unit must promote the use of evaluation results in decision-making and organisational learning by ensuring that policy implications and lessons learnt from (and across) evaluations are synthesised and disseminated. The use of the evaluation results must be regularly monitored by the Evaluation Unit.

²⁹ Communication to the Commission from Ms Grybauskaite in agreement with the President, SEC(2007)213. Unless a case for confidentiality can be made under one of the exceptions provided for in article 4 of Regulation 1048/2001 of the European Parliament and the Council, 30 May 2001. However, evaluations of individual projects financed under programme are not subject to these standards.

³⁰ Council Regulation (EC, Euratom) No 1605/2002 of 25 June 2002.

³¹ http://ec.europa.eu/enlargement/how-does-it-work/financial-assistance/phare/evaluation/interim_en.htm

DG ELARG results-oriented evaluation contributes to DG ELARG's results agenda by ensuring the following arrangements for strictly monitoring the use and follow up of the recommendations:

- ▶ The effective follow up on the recommendations of the evaluation reports is ensured by the Delegations/beneficiaries through the monitoring committee system and special debriefing meetings. Specifically, the outcome of each debriefing meeting with all relevant stakeholders is the agreement on the actions to be taken to address the recommendations of the evaluation report.
- ▶ The interim evaluation scheme involves a systematic follow-up on the findings and recommendations of the sector and country interim evaluation reports. For this purpose, the monitoring committee takes corrective actions taking into account the findings and recommendations of the evaluations.
- ▶ Country summary and thematic evaluation results are followed up immediately after the issue of the report. As a second step, the follow-up of the evaluation recommendations involve that operations reports one year later on what actions have been taken.
- ▶ The QSG helps to verify that evaluation findings are taken into account during the programming phase. It also permits the follow up of lessons learned.
- ▶ Preparation of MIPDs, AOSD reports, and DG ELARG e-manager³² reports represent additional means for follow-up of evaluation results.

Under decentralised management without ex-ante controls, national authorities are responsible to follow-up on their evaluation recommendations and report on them at the JMC.

³² The electronic management reporting tool (e-manager) provides the display of the latest known results in relation to the objectives, including reports on the follow up of the evaluation recommendations.

ANNEX 1: ELARG EVALUATION MANDATE

Mission

The mission of DG ELARG evaluation function is to co-ordinate and implement the evaluation activities of the DG and ensure a proper follow-up of the evaluation findings, results and recommendations. The evaluation function provides inputs for decision-making by contributing to the design of interventions, assisting in an efficient allocation of resources, improving the quality of interventions, and reporting on the accountability of the interventions.

Scope and responsibilities

The scope of evaluation is to make judgements on interventions according to their results, impacts and needs that they aim to satisfy. It shall encompass ex ante, interim and ex post evaluations.

To fulfil its responsibilities, the evaluation function shall:

- Set guidelines for independent and self evaluation;
- Establish the annual evaluation plan;
- Provide the country teams, delegations and beneficiary countries with independent ex ante, interim and ex post evaluation findings;
- Provide a tool for quality management and an early warning system for programme implementation;
- Provide advisory services by drawing out lessons learned and feeding them as recommendations into the design of future programmes;
- Provide for an accountability function of pre-accession assistance;
- Provide an appropriate feedback mechanism to ensure that evaluation results are transmitted effectively to decision-makers;
- Check on how the evaluation results have been taken into account;
- Help develop evaluation capacity in the beneficiary countries;
- Ensure the delivery of and, where appropriate, manage DG ELARG's commitments for evaluation resulting from the Commission's strategic planning and programming cycle and the Communications on evaluation and impact assessment;
- Represent DG ELARG in the external networks in the evaluation field.

Independence and objectivity

The position of the evaluation function in DG ELARG should help integration of the evaluation process into decision-making, ensuring sufficient independence with respect to operational directorates and programme managers. The evaluation function shall ensure that the evaluators' independence in their work is respected and that the evaluation results are not interfered with.

The evaluation function also ensures that the persons carrying out evaluation activities are free of all constraints which could jeopardise their objectivity and behave honourably in their relations with all stakeholders. To maintain objectivity, the evaluation function is not involved in implementing operational programmes and procedures.

Authority

The evaluation function aims to promote the utility and use of evaluation results in the decision-making process. To achieve this, the evaluation function is authorised, in the course of its activities, to:

- Require beneficiary countries and members of staff in the Commission to supply such information and explanations as may be needed within a reasonable period of time;
- Require members of staff and management to report on their use of evaluation results.

Accountability

- The evaluation function shall prepare, in consultation with Operations, a draft annual evaluation plan, including a multi-annual perspective. The evaluation plan shall be consistent with decision-making requirements. Accordingly, it shall ensure that information on the efficiency and effectiveness, and also on the relevance and the utility of policies, programmes and activities is available in good time.
- The draft evaluation plan shall be presented to the Director General for approval.
- The evaluation function shall then co-ordinate and monitor the implementation of this evaluation plan.
- The results of the implementation of the evaluation plan shall be reported to and endorsed by the Director General.
- The evaluation results shall be made publicly available, unless duly substantiated cases of confidentiality.

Standards

The evaluation function shall adhere to the evaluation standards set out in the Communication to the Commission from Ms. Grybauskaite in agreement with the president, Responding to strategic needs: reinforcing the use of evaluation, dated 21 February 2007, SEC (2007)213.

ANNEX 2: MISSION STATEMENT OF ELARG EVALUATION UNIT

The mission of the Evaluation Unit is to co-ordinate and implement the evaluation activities of the DG and ensure a proper follow-up of the evaluation findings, results and recommendations. The Evaluation Unit provides inputs for decision-making by contributing to the design of interventions, assisting in an efficient allocation of resources, improving the quality of interventions, and reporting on the accountability of the interventions.

To fulfil this mission, the Unit shall:

- Set guidelines for independent and self evaluation;
- Help to anticipate decision-making needs;
- Establish a multi-annual evaluation programme and an annual evaluation plan consistent with decision-making requirements;
- Co-ordinate and monitor the implementation of this programme and plan by;
 - serving as a quality management tool for decision-making by drawing out lessons learned and feeding them into the design and implementation of programmes; and
 - providing for an accountability function of pre-accession assistance;
- Provide value for money controls (prospective and retrospective) in order to ensure the economic, efficient and effective achievement of the objectives of the financial assistance;
- Define quality standards and methodology for evaluation activities;
- Communicate effectively the evaluation results to decision-makers and other stakeholders;
- Monitor regularly the use of evaluation results and check on how the evaluation recommendations have been taken into account;
- Help develop evaluation capacity in the beneficiary countries;
- Ensure the delivery of and, where appropriate, manage DG ELARG's commitments for evaluation resulting from the Commission's strategic planning and programming cycle and the Communications on evaluation and impact assessment;
- Represent DG ELARG in the external evaluation networks.

ANNEX 3: FURTHER DESCRIPTION OF VALUE FOR MONEY CONTROLS³³

As defined in the ELARG control strategy document, the Evaluation Unit provides value for money controls (prospective and retrospective), in order to ensure the economic, efficient and effective achievement of the objectives of the financial assistance. The Unit's supervision and control system pays particular attention to contract management matters including detailed work programmes for the external contractors and regular progress reporting, monthly contract management meetings, standard checklist for giving 'conforme aux faits' for invoices, system for quality control of evaluation reports and external independent Quality Assurance Groups.

Value for money controls in centralised (direct and indirect) management

In accordance with the Financial Regulation these value for money controls are applied to all programmes exceeding MEUR 5. The prospective value for money controls provide for judgements on the design and on-going implementation of programmes. The retrospective value for money controls provide for judgements on the achievements of the programmes. The findings of these controls are presented for corrective actions at debriefing and JMC (forthcoming establishment under devolved management) meetings as well as QSG meetings and AOSD reporting.

The selection of the interventions for value for money controls reflects not only the Financial Regulation obligations but also the identification of risks in pre-accession strategic documents, previous evaluations and TOR consultations. The findings of the value for money controls of evaluations are reported at sectoral, country, thematic and consolidated programme level.

Value for money controls in decentralised management with ex-ante control

The value for money controls (prospective and retrospective) of DIS supports the Director General, AOSDs and the national authorities. In accordance with the JMC mandate all monitoring sectors are evaluated once per annum and the evaluation findings are presented for corrective action through AOSD reporting and at the JMC and debriefing meetings.

The findings of the prospective value for money controls (mainly ex-ante evaluation of MIPDs)/annual programmes) are presented as recommendations and lessons learned at the deliberations of the QSG and through AOSD reporting.

Value for money controls in decentralised management without ex-ante control

In accordance with the JMC mandate for the New Member States all monitoring sectors shall be subject to an Interim Evaluation. The management of the IE evaluation is decentralised to the National Authorities at the same time as ex-ante controls are waived. However, the responsibility for ex-post evaluations remains within the DG ELARG Evaluation Unit.

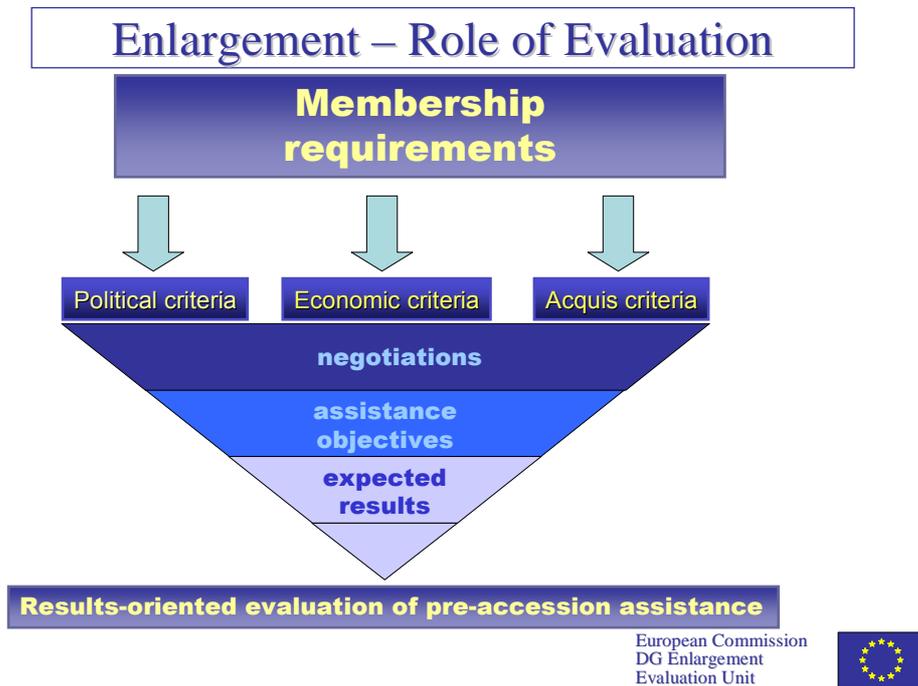
The findings on these value for money controls are reported and followed up at the JMC meetings and in the Implementation Status Reports by the National Authorities as well as in the AOSD reporting.

³³ DG ELARG Control Strategy in 2008, section 4.6. Further Description of Value for Money Controls.

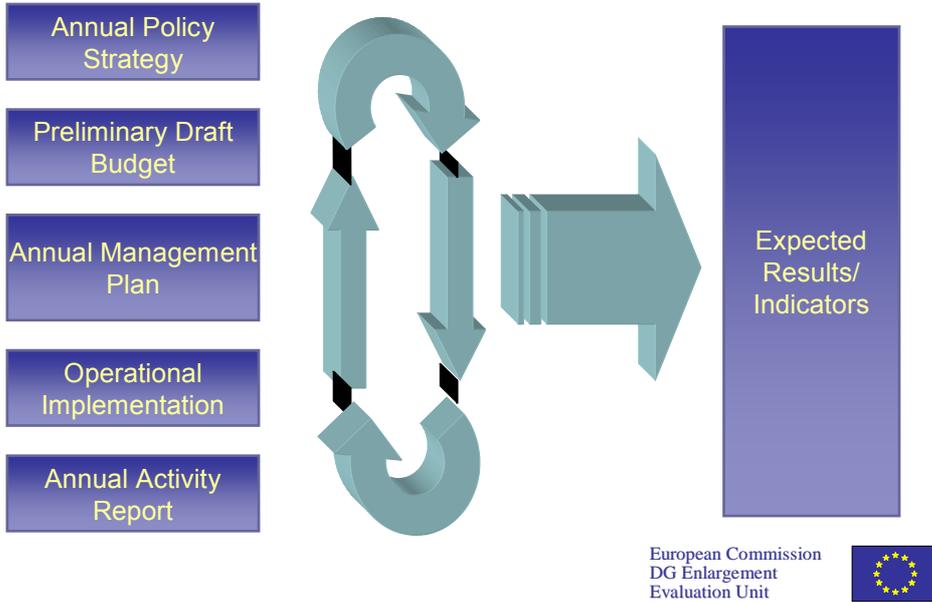
Value for money controls in joint management

Value for money controls in joint management (SIGMA and SMEFF) are undertaken from time to time.

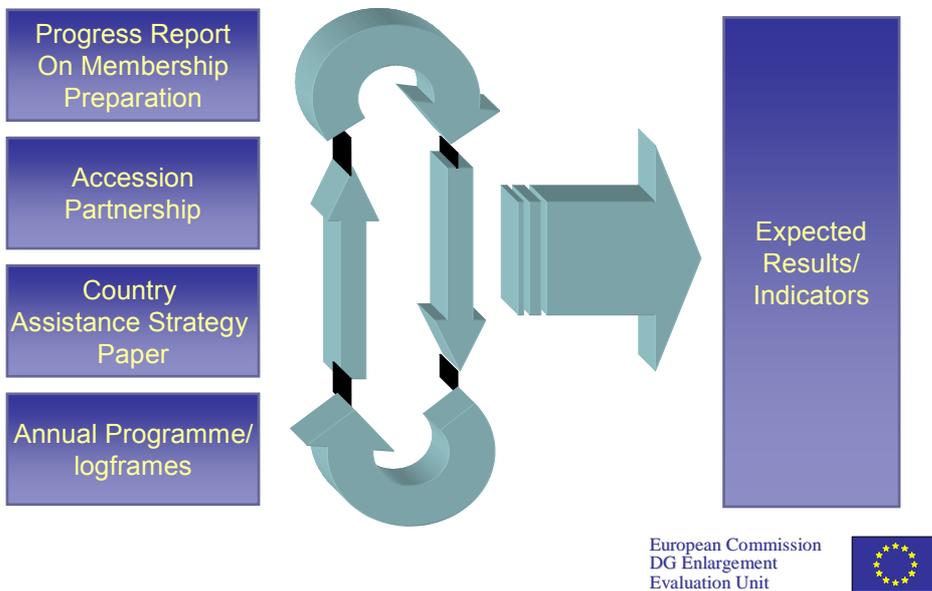
ANNEX 4: PRESENTATION ON DG ELARG EVALUATION FUNCTION



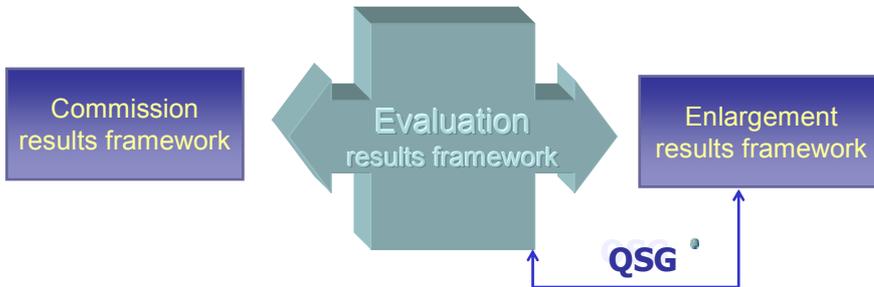
Commission Multi-annual Results Framework



Enlargement Results Framework



Linking Results Frameworks

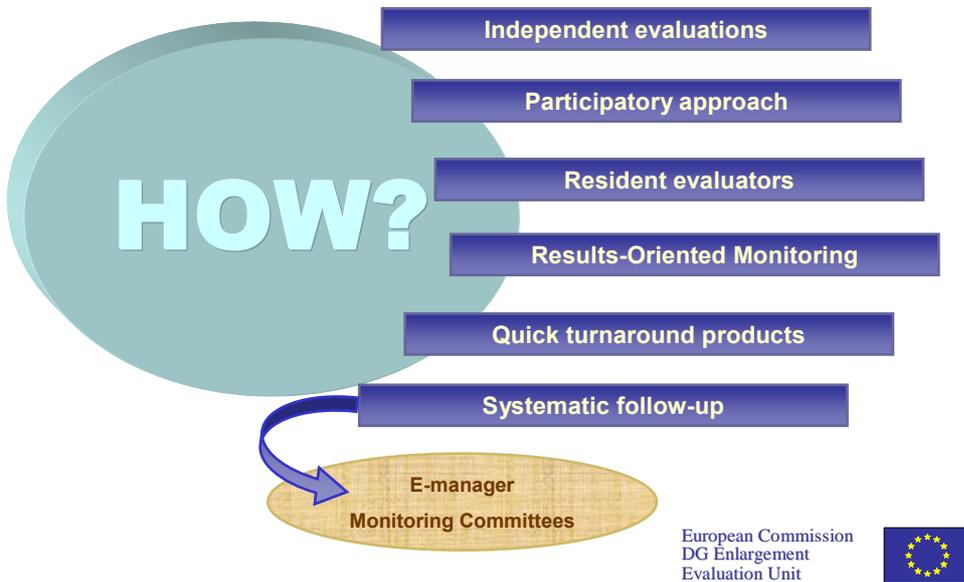


Quality Support Group 

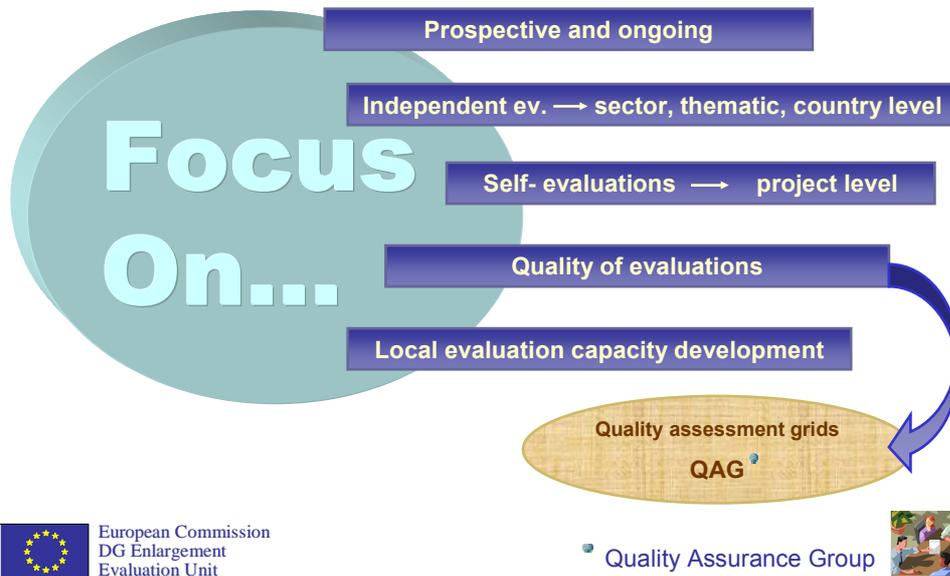
European Commission
DG Enlargement
Evaluation Unit



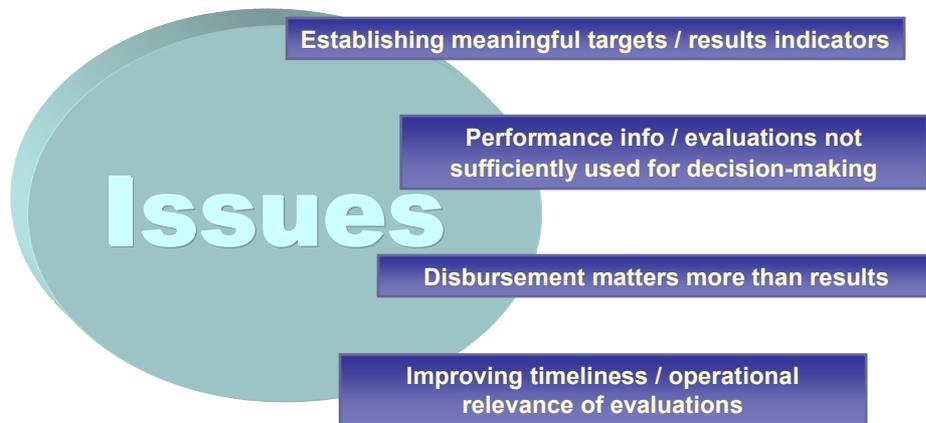
Evaluation Results Framework (1)



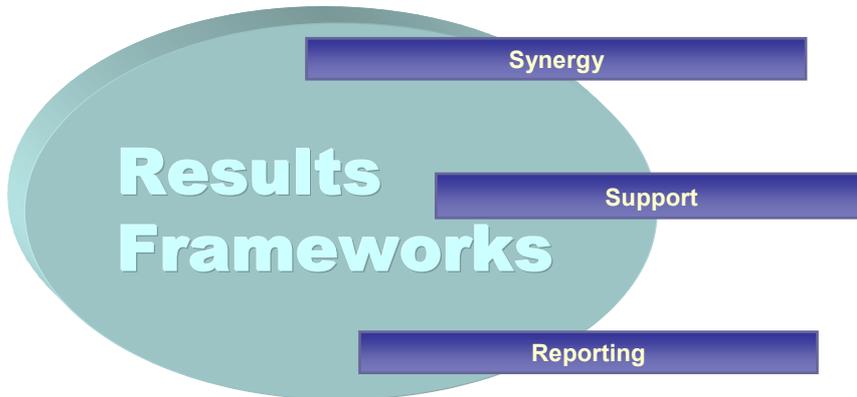
Evaluation Results Framework (2)



Challenges



Gradual and Steady Progress

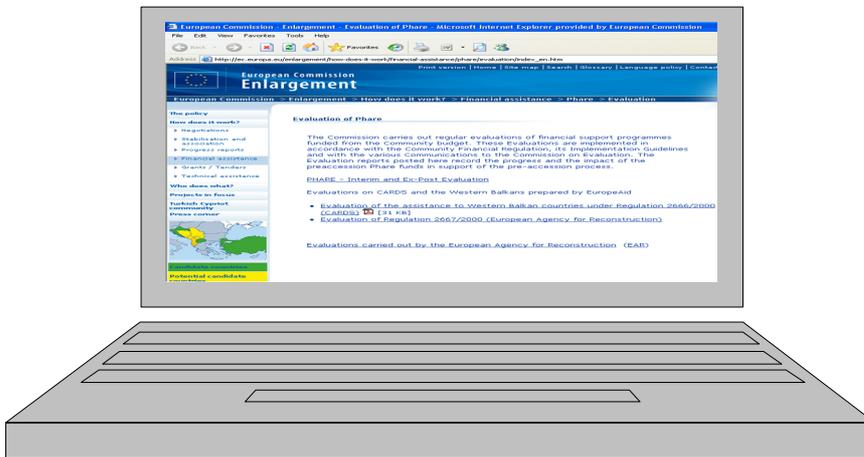


European Commission
DG Enlargement
Evaluation Unit



For Further Information

http://ec.europa.eu/enlargement/how-does-it-work/financial-assistance/phare/evaluation/index_en.htm



European Commission
DG Enlargement
Evaluation Unit



ANNEX 5: EXAMPLE OF TERMS OF REFERENCE

Phare *Ex Post* Evaluation

Phase 1. Multi-Beneficiary Programmes

Small and Medium-Sized Enterprise Finance Facility (SMEFF)

Terms of Reference

OBJECTIVES

The purpose of this evaluation is to assess the contribution of the Small and Medium-Sized Enterprise Finance Facility (SMEFF) in inducing financial intermediaries to expand and maintain, in the long term, their financing operations with Small and Medium Sized Enterprises (SMEs).

The evaluation of the SMEFF MBP (Multi Beneficiary Programme) is one of a series of MBP evaluations in different areas. It will become a stand alone report but will also feed into a consolidated evaluation of Phare multi-beneficiary programmes. The evaluation of the SMEFF will also feed into the section in the thematic evaluation of the SME sector that deals with issues of finance. Both of these will in turn form part of a consolidated ex-post evaluation of the Phare programme.

BACKGROUND AND CONTEXT

The specific objectives of the SMEFF have aimed at inducing financial intermediaries to expand and maintain their long term financing operations with SMEs. These objectives have, since the start of the facility in 1999, evolved over time and in the year 2003 comprise of:

- Building further the confidence of financial intermediaries to engage in financial operations with SMEs;
- Enhancing further the capacity of financial intermediaries to assess and monitor the related risks and to manage their financial exposure;
- Deepening country coverage at a regional level by concluding more projects within each country; and,
- Sustaining and maximising the impact of existing Projects under Loan, Guarantee and Leasing Window by providing credit lines to the financial intermediaries concerned.

During the ex-post evaluation period (1996 – 2001), the SMEFF assistance was provided by six MBPs³⁴ with a total value of € 191 million. During the period of interim evaluation 2002 and 2004, a further M€ 100 was provided to the SMEFF under another six MBPs.³⁵ The ongoing post-2001 programmes will also be evaluated. The value of the *ex post* evaluation will be extended in this way by checking to what extent lessons learned up to 2001 have been incorporated in the subsequent programmes. Moreover, the interim evaluation will ensure that the evaluation results are more up to date as inputs for decision making.

Support to SMEs was also provided under Phare national programmes, as well as by other donors, such as World Bank, UNDP etc. The SMEFF is primarily designed to support financial institutions.

The EMS Interim Evaluation (IE) Report on the SMEFF, finalised in February 2003³⁶, concluded that while the SMEFF was a good concept, with a high potential for success, it required fine-tuning across a range of areas to respect Phare Guidelines related to catalytic effect, additionality and sustainability. Key issues identified in the IE Report included:

- *Ensuring catalytic impact by identifying and addressing market failures*

The IE concluded that there was scope for the SMEFF as an instrument to be more responsive to regional (both inter and intra country) differences, differing market segments (including the less attractive market and product segments such as micro-financing and start-ups) and in adapting to meet changing needs to ensure its continued relevance. The report recommended a needs assessment of SME financing including identifying differences in market maturity, at country and regional level and in market/product segments that experience particular difficulty in accessing finance.

- *Ensuring additionality*

In most cases the Participating Banks (PBs) are already present in the SME market and will remain active. Moreover, in several cases the PBs have major active foreign shareholders. Much more attention should therefore be paid to the selection of Banks and of projects to ensure that already strong and experienced PBs are not displacing other financial intermediaries such as smaller, regional locally owned banks.

- *Ensure sustainability by focussing on capacity building at Financial Institutions*

A core anticipated impact of the SMEFF is the development of a successful business model for addressing the SME market, underpinned at a strategic level within the PBs. The IE report highlighted the weakness of the institution building components of the SMEFF and this will be an important focus for this *ex post* evaluation, particularly for remaining funds under the SMEFF in both the new member states and the remaining candidate countries (CCs).

Capacity building at financial institutions through the delivery of technical assistance is an important part of the SMEFF. However, the IE report concluded that the technical assistance had not delivered the expected value for money for different reasons: the commitment of the management of the PB, the 'fit' between the consultants and the

34 Projects 1999/001-979, 2000/002-046, 2000/002-052, 2001/002-077, 2001/002-079, and 2001/002-095.

35 Projects 2002/000-621, 2002/000-622, 2002/000-629, 2003/005-745, 2003/005-765, and 2003/005-766.

36 R/ZZ/SME/02.146 issued on 18 February 2003.

needs of the PB, and the management by the International Financing Institutions (IFIs). This has had an adverse impact on sustainability.

- *Clarify the role of, and approach to, the Equity Window*

In the EU-15 member states, equity funds open to smaller businesses are considered a useful instrument to support SME development. In the new member states and remaining CCs however, the availability of such funding is relatively low. The IE report found that the implementation of the Equity Window (EW) component of the SMEFF to be proceeding at a slow pace³⁷. The demonstration effect is largely dependant on the funds achieving profit in the longer term, which will in turn motivate other providers of SME equity finance to enter the market. Delays in implementation will therefore have an adverse effect on potential impact. Furthermore, even operating at its optimum, the number of SMEs that could benefit from the EW Fund would be a maximum of around 100 investments, thus confirming its effect as a demonstration, rather than a deep impact.

- *Operational issues*

Operational issues identified included : a) the need to ensure a uniform quality of reporting from the participating IFIs to the Commission Services and; b) more transparent management of the PB pipelines, including the availability of information on terms and conditions offered by all participating IFIs to the PBs and leasing affiliates.

EVALUATION QUESTIONS

The evaluation questions and related judgement criteria for this exercise will be derived from:

- Performance evaluation questions
- Thematic/cross-cutting questions

Performance evaluation questions

Needs assessment and design

The evaluation will start with an assessment of whether the relevance and design of the SMEFF was coherent with the principles of catalytic impact, additionality and sustainability. In particular, attention will be paid to whether the analysis of the demand for SME finance, took sufficiently into account known and/or potential needs, e.g. level of development of the local market, market failures, need for institution building, type of capital required (equity versus debt), presence of foreign-owned banks, and access to venture capital.

The extent to which inputs/activities have produced outputs/results

The outputs envisaged by the SMEFF are: i) expanded capacities of the PBs and leasing affiliates to serve the needs of SMEs and, iii) a demonstration effect to other

³⁷ R/ZZ/SME/02.146 issued on 18 February 2003.

potential financial intermediaries that operating with SMEs is viable; ii) greater access to finance among participating SMEs in the region.

The evaluation will assess the cost effectiveness of the activities. The evaluation will review whether the technical assistance provided to the PBs and leasing affiliates was appropriate.

The extent to which the results/impacts contributed to achieving wider objectives

The specific objectives have been aimed at inducing financial intermediaries to expand and maintain in the long term their financing operations with SMEs. The evaluation will assess the extent to which the specific objectives of building further the confidence of financial intermediaries to engage in financial operations with SMEs, enhancing further the capacity of financial intermediaries to assess and monitor the related risks and to manage their financial exposure, deepening country coverage at a regional level by concluding more projects within each country and sustaining and maximising the impact of existing Projects under Loan, Guarantee and Leasing Window by providing credit lines to the financial intermediaries concerned have been achieved.

Long-term viability following the withdrawal of Phare support

Long-term sustainability will be investigated at a number of levels. Linked to the overall objective, the focus will be on whether a successful business model has been developed in the PBs/leasing affiliates, directly or indirectly attributable to their participation in the SMEFF. This will include for example, whether PBs/leasing affiliates have adopted new strategies, policies, procedures and attitudes to undertaking, marketing actions, managing and pricing SME credit risk.

Thematic/cross-cutting questions

The extent to which Phare support improved the performance of CC beneficiaries

The purpose of this question is to assess whether Phare support through the SMEFF in practice addressed the *ex post* needs of the beneficiaries. Thus the evaluation would seek to assess whether the original objectives (*ex ante* needs) of the programme have evolved appropriately. This assessment would augment the analysis based on the five performance evaluation criteria.

Complementarity of SMEFF to other SME financial support measures

The purpose of this question is to locate the role of the SMEFF within the context of other financial support measures to stimulate development of the SME sector such as SME support under economic and social cohesion.

Effectiveness and efficiency of the MBPs as a delivery mechanism

There are inherent strengths and weaknesses with centralised MBPs. Positive factors include economies of scale in design and implementation, scope for more effective cross-fertilisation, and the ability to utilise the expertise of the specialist bodies to which management is delegated. Negative factors include difficulties in effective monitoring by

the Commission and the risks associated with a “one size fits all” approach. With particular reference to the SMEFF, the *ex post* evaluation will consider the extent to which the multi-beneficiary approach has delivered the required results.

METHODOLOGY

To reflect the high level of interest and strategic value of the SMEFF MBP, this evaluation will be in-depth, and will include not only a strong basis of retrospective assessment (the *ex post* evaluation of the 1996 – 2001 programmes) but also an interim evaluation of post-2001 allocations.

Following an initial collection of available documents, the evaluation will start with a desk review. This will be followed by further data gathering using the most appropriate tools (interviews or focus groups in Brussels and/or the CCs, and/or questionnaires)³⁸. The evaluation will use a sample of participating banks/SMEs selected based on:

- The distribution of funds between the sponsoring organisations (EBRD, EIB, CEB/KfW);
- The distribution of funds between participating banks;
- The geographical focus – including representative samples of countries focussing on four new member states [Poland, Hungary, Slovakia and Estonia], plus Bulgaria and Romania.

To support the evaluation questions a set of judgement criteria and evaluation indicators will be developed. These may be both quantitative and/or qualitative. A preliminary set of indicators will be discussed during the kick-off process with the aim of ensuring that requests for information relating to the indicators are understood in context. The preparation of a set of indicators will take into account the indicators of achievement related to the Institution building component that has recently been introduced by the sponsors.

Representatives of stakeholders and target audiences will be invited to participate during the kick-off process which may involve a meeting and/or electronic consultation meeting. This process is crucial for ensuring a common understanding of the evaluation questions and related indicators, and for allowing maximum participation. Stakeholders will also be invited to join a virtual steering group.

REPORTING AND TARGET AUDIENCES

The main users of the evaluation will be DG ELARG Directorate responsible for the SMEFF, DG ECFIN, and DG Enterprise, EC Delegations/Representations, and the National Aid Co-ordinators of beneficiary countries. In addition, country teams for the Western Balkans and Turkey will be important users of the evaluation results.

³⁸ Interviews and questionnaires may be carried out by telephone, if appropriate.

ACTIVITIES, RESOURCES AND TIMETABLE

The SMEFF MBP evaluation will be conducted in a number of stages as follows:

Step	Activity	2005				2006	
		Sep	Oct	Nov	Dec	Jan	Feb
1	Preparation	■	■	■			
2	Info gathering and processing			■	■		
3	Drafting of report for E4				■		
4	Comments from E4					■	
5	Drafting of final SMEFF MBP report and submission to Consolidated MBP Report					■	

The evaluation will be carried out by a team consisting of the deputy project director, a key expert, other senior and junior experts, and short-term technical specialists (SSTS), both international and local. Since the SMEFF exercise is an in-depth evaluation, the total resource envelope available for this exercise amounts to 90 man-days.

ANNEX 6: EVALUATION CRITERIA

The basic concepts and the evaluation methodology applied to the various types of evaluation are based on key evaluation criteria: relevance, efficiency, effectiveness, impact and sustainability. The methodology is also consistent with DG Relex/Aidco approach and reflects the elements and concepts of intervention logic that analyses a project along its inputs, outputs and effects.

1. Relevance-Main message

The relevance of a project relates primarily to its design and concerns the extent to which **its stated objectives correctly address the identified problems or real needs**. It needs to be kept under review throughout the life of the project in case changes occur either in the nature of the very problems originally identified, or in the circumstances – whether physical, political, economic, social, environmental, institutional or policy - in which the project takes place, necessitating a corresponding change of focus.

In other words, **relevance concerns the appropriateness of the project design to the problems to be resolved at two points in time: when the project was designed, and at the time of the evaluation**. However, the flexible *handling* of any changes needed to keep a project relevant forms part of the definitions of the other evaluation criteria, given below.

In an evaluation, an analysis of Relevance will therefore systematically focus on the following:

- **identification of real** (as distinct from perceived) **problems or needs**, and of the **correct beneficiaries**, and how well the project's initial design addressed them;
- quality of assessment of **local absorption capacities**;
- quality of assessment of **local implementation capacities**;
- **preparatory activities** undertaken (policy assessments, sector reviews, [pre-]feasibility studies including financial and economic analysis, planning workshops, etc), by whom, how well the findings were incorporated into the final project document, and any obvious omissions;
- appropriateness of **initial consultations with**, and **participation by**, local key stakeholders including the Delegation, national authorities, intended beneficiaries, and other donors (the last-mentioned especially on complementarity aspects) before the design was confirmed and implementation started;
- **complementarity and coherence with related activities undertaken elsewhere** by government or other donors, at the same level or at a higher level, rather than duplication or conflict;
- the quality of the entries in the **assumptions, risks and conditions** column of the LogFrame at the appropriate levels;

- overall **design strengths and weaknesses** including:
- quality of the **LogFrame** (or LogFrames if a multi-component programme);
- clarity and internal consistency of the stated **overall objectives, purpose and results**;
- whether the objectively-verifiable **indicators** of achievement (OVIs) were well-chosen and widely agreed;
- realism in choice and quantity of **inputs**;
- overall **degree of flexibility** and adaptability to facilitate rapid responses to changes in circumstances.

2. Efficiency

The efficiency criterion concerns how well the various **activities** have transformed the available resources into the intended **results** (sometimes referred to as *outputs*), in terms of quantity, quality and timeliness. A key question it asks is “**have things been done right?**” and thereby also addresses value-for-money, e.g., whether similar results could have been achieved better by other means at lower cost within the same amount of time.

An analysis of Efficiency will therefore focus on:

- **the quality of day-to-day management**, for example in:
 - management of the budget (including whether an inadequate budget was a factor);
 - management of personnel, information, property, etc;
 - whether management of **risk** was adequate, i.e. whether flexibility was demonstrated in response to changes in circumstances;
 - relations/co-ordination with local authorities, institutions, beneficiaries, other donors;
 - respect for deadlines;
- **costs and value-for-money**: how far the costs of the project were justified by the benefits - whether or not expressed in monetary terms - that they generated³⁹, in comparison with similar projects or known alternative approaches, taking account of contextual differences;
- **candidate country contributions** from local institutions and government (e.g. offices, experts, reports, tax exemption, as set out in the LogFrame resource schedule), target beneficiaries and other local parties: were they provided as

³⁹ A comparison broadly known as cost-effectiveness analysis (similar to cost benefit analysis except that the latter applies where benefits are measurable in monetary terms).

planned, could reallocation of responsibilities have improved performance, has communication been good?

- **Commission HQ/Delegation inputs** (e.g. procurement, training, contracting, either direct or *via* consultants/bureaux): key questions as for local/government inputs (above);
- **technical assistance**: how well did it help to provide appropriate solutions and develop local capacities to define and produce results?
- **quality of monitoring**: its existence (or not), accuracy, flexibility and utility; adequacy of baseline information;
- whether the chosen **indicators** of efficiency were suitable and, if not, whether management amended them;
- did any **unplanned results** arise from the activities?

3. Effectiveness

The effectiveness criterion, in the LogFrame terminology, concerns how far the project's **results** were *used* or their potential benefits were realised - in other words, whether they achieved the project **purpose**. The key question is what difference the project made in practice, as measured by how far the intended beneficiaries really benefited from the products or services it made available.

The analysis of Effectiveness will therefore focus on:

- **whether the planned benefits have been delivered and received**, as perceived mainly by the key beneficiaries, but also taking account of the views of donor management, the responsible national Government authorities, and other concerned parties (NGOs, business associations etc);
- the appropriateness of the **indicators of benefit** used in the above assessment to measure achievement of the project purpose (this is also relevant to cost-effectiveness analysis as referred to in the footnote on the previous page); this should include a judgement on how promptly and effectively the project management reacted to any changes that occurred following the initial design by amending indicators found no longer to be appropriate;
- in institutional reform projects, **whether behavioural patterns have changed** in the beneficiary organisations or groups at various levels; and how far the changed institutional arrangements and characteristics have produced the planned improvements (e.g. in communications, productivity, ability to generate actions which lead to economic and social development);
- if the **assumptions and risk assessments** at results level turned out to be inadequate or invalid, or **unforeseen external factors** intervened, **how flexibly management adapted** to ensure that the results would still achieve the purpose; and how well it was supported in this by key stakeholders including Government,

Commission (HQ and locally), etc.: in summary, “**were the right things done**” to ensure that the potential beneficiaries actually benefited?

- whether the **balance of responsibilities** between the various stakeholders was correct, which accompanying measures were or should have been taken by the partner authorities, and with what consequences;
- how **unplanned results** may have affected the benefits received;
- whether any shortcomings at this level were due to a failure to take account of **crosscutting or over-arching issues** such as gender, environment and poverty during implementation.

4. Impact

The term impact, sometimes referred to as *outcome*, denotes **the relationship between the project’s purpose and overall objectives**, that is the extent to which the benefits received by the target beneficiaries had a wider overall effect on larger numbers of people in the sector or region or in the country as a whole.

The prospects for *Impact* and *Sustainability* are harder to assess than for the previous three criteria. They are subject to factors which are not within the project’s control. So, rating these criteria requires the evaluator to look beyond the confines of the project.

Moreover, while there may on occasion be some impacts - contributions to the overall objective of the project -during its lifetime, actual impact and, by definition, actual sustainability can only be fully assessed after - sometimes well after - the project has ended.

It follows that assessment of impact and sustainability may be particularly difficult when the project in question is at an early stage of implementation.

Evaluators should therefore examine and rate likely impact on the basis of the extent to which the necessary pre-conditions for impact are in place, or are being put in place, using a checklist. It should be noted that some pre-conditions are common to both impact and sustainability.

A model checklist is given below. However, the detailed issues to be taken into account will need to be considered on a case by case basis. They will vary with the sector, country, institutional set-up etc, and are for the evaluator to determine, and embody in a definitive checklist.

Checklist:

Are the following pre-conditions for **Impact** in place or being put in place?

- Is the project conceived within a strategy framework?
- Is it based on a needs assessment?

- Are the essential needs for adequate public and administrative capacity in place, including stable institutions, appropriate public investment policies and adequate inter-ministerial and central-regional collaborative machinery of government?
- Have you rated effectiveness in achieving project's objectives positively?
- Will the project's outputs contribute to the overall objective?
- Will there be an identifiable benefit for society or the economy?
- Are all the public administrative units / Ministries / regional structures involved in the ongoing implementation of the activity supported aware of it? Do they have complementary public administrative systems in place?
- Have the ultimate recipients of / customers for / the outcomes of the project been made aware of it?
- Have relevant civil society bodies been involved in the project and are their roles defined in relation to its outputs?
- Are there other considerations relevant to the particular project?

5. Sustainability

The fifth and final - and often most important - criterion, sustainability, relates to **whether the positive outcomes of the project at purpose level are likely to continue after external funding ends**, and also whether its longer-term impact on the wider development process can also be sustained at the level of the sector, region or country.

As stated above (para. 4), the prospects for Sustainability may be difficult to assess, given that they are subject to factors outside the project's control. In particular, assessment is hard when a project is at an early stage of implementation.

Evaluators should therefore examine and rate likely sustainability on the basis of the extent to which the necessary pre-conditions for sustainability are in place, or are being put in place, using a checklist. It should be noted that some pre-conditions are common to both impact and sustainability.

A model checklist is given below. However, the detailed issues to be taken into account will need to be considered on a case by case basis. They will vary with the sector, country, institutional set-up etc, and are for the evaluator to determine, and embody in a definitive checklist.

Checklist:

Are the following pre-conditions for **Sustainability** in place or being put in place?

- Have you rated effectiveness in achieving project's objectives positively?
- Is there a sectoral strategy document within which the need for the project was identified and to which the outputs will contribute?

- Do those with strategy / policy / management responsibility for ongoing implementation of the outputs demonstrate 'ownership'?
- Are the horizontal public administration systems stable and adequate?
- Is there ongoing national finance available, including for maintenance, replacement, insurance, disposables?
- Are the provisions for ongoing staffing / staff replacement / training secure?
- Are procedures and systems fully documented, with defined responsibility for updating?
- Are there other considerations relevant to the particular project?

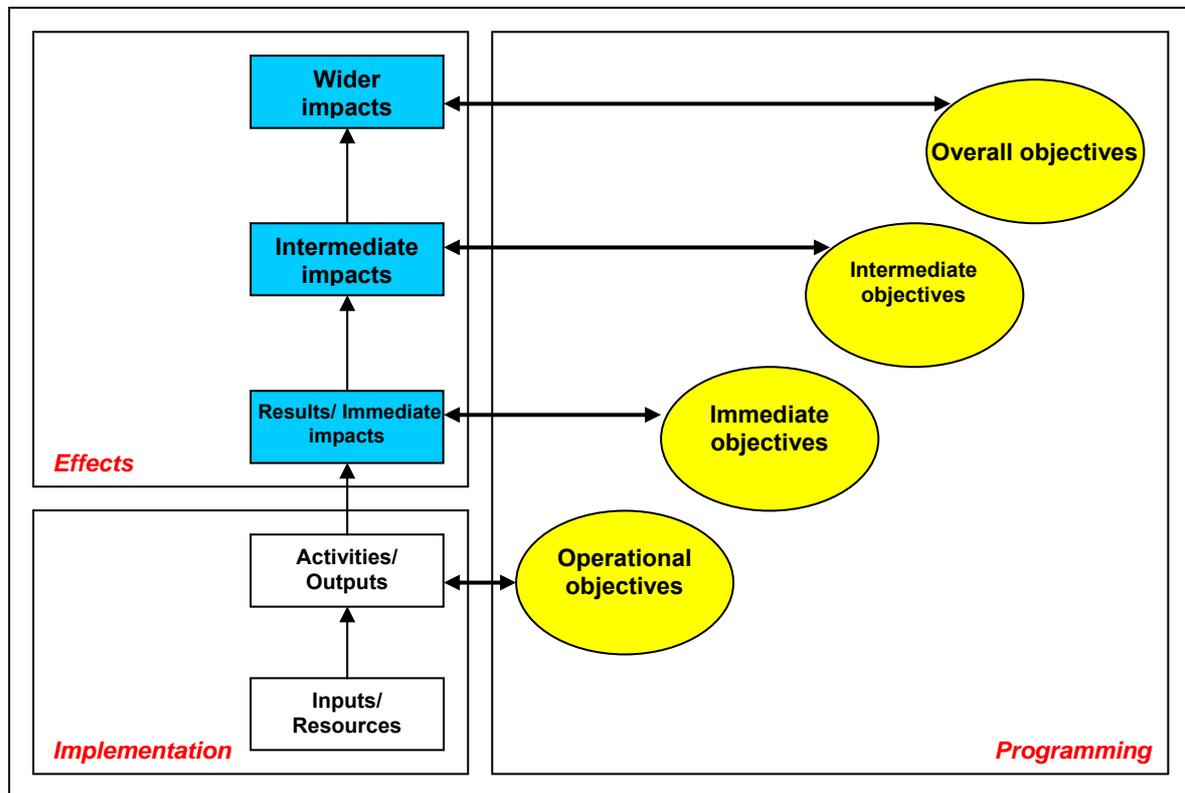
Annex 7 elaborates further on the five evaluation criteria in relation with the intervention logic.

ANNEX 7: MORE ON THE FIVE EVALUATION CRITERIA AND THE ELEMENTS OF INTERVENTION LOGIC

The following charts set out the main linkages between the elements of intervention logic and respectively the hierarchy of project objectives (chart 1), and the key evaluation criteria, including immediate, intermediate and wider impacts (chart 2).

Chart 1: Links between the LogFrame elements and the hierarchy of project objectives

The elements and concepts of intervention logic as presented below form, by definition, the core of evaluative inquiry, they do not represent an exhaustive list and additional issues, where relevant, may also be examined within the framework of an evaluation. A framework for developing a hierarchy of objectives and associated indicators is presented below.



The fixing of objectives when an intervention is being designed is an essential aid to evaluation since they are an integral part of assessing an intervention directly with regard to the issues of *relevance* and *effectiveness*. Furthermore, objectives provide a starting point from which indicators for measuring performance (through both monitoring mechanisms and evaluations) should be developed.

Operational objectives provide a basis for assessing an intervention in relation to its *outputs*. The latter can be defined as what is directly produced/ supplied through the implementation process. Indicators at this level are called output indicators.

Specific objectives provide a basis for assessing an intervention in relation to the short-term *results* that occur at the level of direct beneficiaries/recipients of assistance. Indicators at this level are called results indicators.

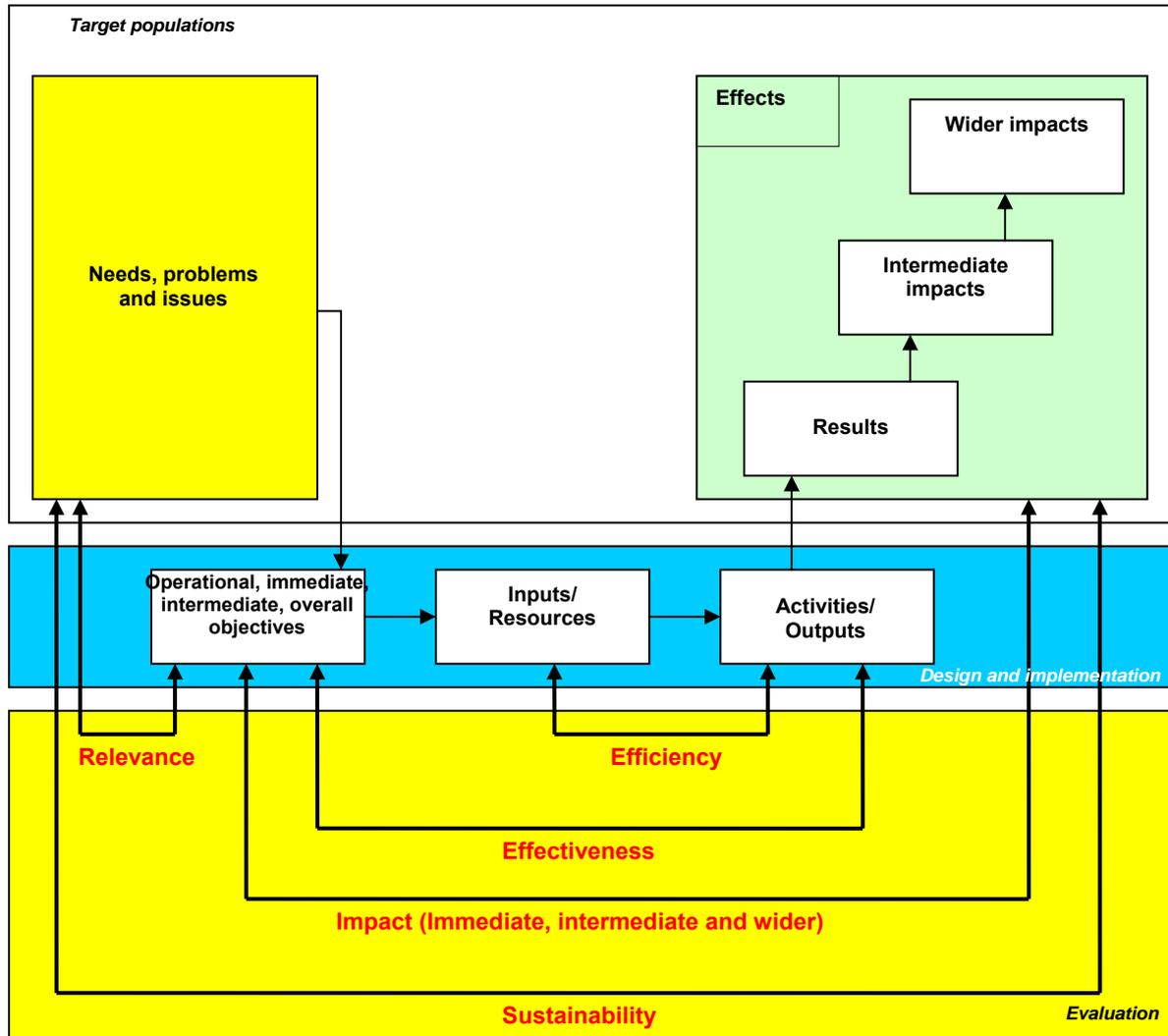
Intermediate objectives provide a basis for assessing an intervention in relation to its short to medium-term effects on both direct and indirect beneficiaries/recipients of assistance. Indicators at this level are called impact indicators.

Global objectives provide a basis for assessing an intervention in relation to longer term and more diffuse effects (or *global impacts*). Indicators at this level are also called impact indicators.

Inputs: these are the means used to produce outputs. Inputs include budgetary costs (financial, administrative and human resources), but also costs for the beneficiaries or target population (co-financing, compliance costs stemming from administrative burden).

Outputs: Output is defined as a product, which is delivered by the project. It is typically a product, which is under direct control of the manager. When specifying output, you need to ask i.) what should be delivered, and ii.) at what time.

Chart 2: Links between the elements of intervention logic and the evaluation criteria



Within this framework, the way in which an intervention is designed and implemented is considered as an explanatory variable, along with elements of the “external” environment (competing or concurrent factors) and the effects of other public interventions.

ANNEX 8: RATING FOR QUALITY ASSESSMENT

A. GUIDE ON SCORING THE CRITERIA

1. RELEVANCE

Does the evaluation respond to information needs, in particular as expressed in the terms of references?

This criterion concerns how well the evaluation responds to the terms of references.

- The evaluation report deals with and responds to the evaluation questions
- A justification was provided if any evaluation questions wasn't answered
- The scope covers the requested periods of time, geographical areas, target groups, parts of budget, regulations, etc
- Limitations in scope are discussed and justified
- Effects on other policies, programs, groups, areas etc are considered
- Unintended effects are identified
- The evolution of the program is taken into account possible changes in the problems and needs compared to the situation at the start of the intervention have been addressed
- The evaluation broaden the scope or enlighten the approaches in the policy cycle
- The evaluation add value to existing policy knowledge
- Other

2. APPROPRIATE DESIGN

Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?

This criterion concerns the inception phase. It operationalises and possibly complements the terms of reference. In some cases, because of unforeseen events, it may also relate to a subsequent reorientation of parts of the evaluation work.

- The rationale of the program, cause-effect relations, outcomes, policy context, stakeholder interests, etc have been studied and taken into account in design the evaluation
- The evaluation methodology chosen is coherent with evaluation needs and requests
- The methodology is clearly and adequately described, in enough detail for the quality to be judged. It is described to the extent that the evaluation can be

<p>replicated</p> <ul style="list-style-type: none"> • Information sources and analysis tools are adequate for answering the evaluation questions • Criteria for judgement are anticipated • Weaknesses of the selected methodology are pointed out along with potential risks • Other methodological alternatives are considered; their pros and cons are explained • Research design has been validated with experts or relevant stakeholders if appropriate (e.g. experts on related policies, specific evaluation know-how) • Ethical issues are properly considered (confidentiality of sources of information, potential harms or difficulties of participation of stakeholders, etc) • Other
--

3. RELIABLE DATA

Are data collected adequate for their intended use and have their reliability been ascertained?

This criterion concerns the relevance and correctness of both primary and secondary data.

- Available information and sources are well identified
- Relevant literature and previous studies have been sufficiently reviewed
- Existing monitoring systems were used
- Data and information are free from factual or logic errors; data gathered are correct and sufficient
- Data collection rationale is explained; and it is coherent with the design of the study
- The quality of existing or collected data was checked and ascertained
- The amount of qualitative information and quantitative data is balanced and appropriate for a valid and reliable analysis
- Tools and means used to collect and process data (e.g. surveys, case studies, expert groups, etc...) were: selected in relation to established criteria; complete and suitable for answering the evaluative questions; adequately used as to guarantee the reliability and validity of results
- Tools and data collection limitations (missing coverage, non-participation or non-attendance of selected cases) are discussed and explained.
- Correcting measures have been taken to avoid any potential bias and/or their implications
- Other

4. SOUND ANALYSIS

Is information systematically analysed to answer evaluation questions in a valid manner?

This criterion refers to the correct interpretation of data and to the adequacy of the methodology applied.

- There is a clear, solid and coherent deductive analysis (e.g. controlled comparison, experimental research, inferential statistics, etc...)
- The analysis is well focussed on the most relevant cause/effect relations and influences underlying the program logic
- The analysis uses appropriate quantitative or qualitative techniques, suitable to the evaluation context
- Cross checking of findings has taken place. The analysis relies on two or more independent lines of evidence
- Explanatory arguments are explicitly (or implicitly) presented
- The context (historical, socio-economic, etc...) is well taken into account in the analysis
- The report reflects an appropriate range of stakeholders consulted
- Inputs from important stakeholders are used in a balanced way
- The limitations of the analysis and exceptions to general explanations or evidences were identified, discussed and transparently presented
- Other

5. CREDIBLE FINDINGS

Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?

This criterion concerns the coherence of the findings with the preceding analysis and data.

- Judgements are based on transparent criteria
- Findings are supported by evidence originating from sound analysis
- Generalisations or extrapolations, when made, are justified (e.g., through the sampling or selection of cases)
- Findings corroborate existing knowledge; differences or contradictions with existing know-how are explained
- Stakeholder opinions were taken into account for interpreting empirical facts
- Main findings are testable
- Limitations on validity are pointed out; trade-offs between internal and external

<p>validity are identified and discussed</p> <ul style="list-style-type: none"> • Results of the analysis reflect an acceptable compromise of the perceptions of stakeholders and those described by figures and facts observed and estimated • Other

6. VALID CONCLUSIONS

<p><i>Are conclusions non-bias and fully based on findings?</i></p>
<p>This criterion concerns the extent to which conclusions logically stem from findings and are based on impartial judgement criteria.</p> <ul style="list-style-type: none"> • Conclusions are properly addressed to the evaluations questions • Conclusions are coherently and logically substantiated by evaluation findings • There are no relevant conclusions missing according to the evidences presented • Conclusions are interpreted in relation with the policy context • Conclusions are free of personal or partisan considerations; potential influence of values and interests of the evaluation team in the research methodology and outcome are openly discussed • Conclusions are orderly presented and related (categorised, ranked, priorities, sequence) • Controversial issues are presented in a fair and balanced manner • Other

7. HELPFUL RECOMMENDATIONS

<p><i>Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic?</i></p>
<p>This criterion concerns the soundness and realism of the recommendations</p> <ul style="list-style-type: none"> • Recommendations stem logically from conclusions • Plausible options for improvements are identified • Recommendations covers all relevant main conclusions • They are realistic and potentially useful • Relations among recommendations are taken into account (e.g. priority ranking, sequencing, etc) • Recommendations provide certain guidance for action planning • Where feasible, cost of recommendations were estimated • Other

8. EXECUTIVE SUMMARY

Does the executive summary provide a synthesis of the main findings of the report and bring out the key messages of it?

This criterion concerns a well structured, stand-alone document of key messages.

- The report must have an executive summary presented in the form of a synthesis
- It should be short, no more than five pages
- The report should be a tightly drafted, to the point stand-alone document
- It must bring out all the key issues/analytical points
- It must bring out the key messages/specific recommendations
- Key recommendations must be carefully targeted to the appropriate audiences at all levels

9. CLARITY

Does the report adequately describe the policy being evaluated and its context? Is it well structured and written in an understandable manner?

This criterion concerns to the clarity of the presentation and the appropriateness of the content of the evaluation.

- The content of the report describes the policy being evaluated, its context, the evaluation purposes, contextual limitations, methodology, findings, etc in a neat and well structured manner
- The report is well structured and signposted to guide and facilitate reading
- Key messages are summarised and highlighted
- There is a clear presentational linked sequence among data, interpretation and conclusions
- The report includes a relevant and concise executive summary, which includes main conclusions and recommendations in a balanced and impartial manner
- Specialised concepts were used only when necessary and were they clearly defined
- Tables, graphs, and similar presentational tools are used to facilitate understanding; they are well commented with narrative text
- The length of the report (excluded appendices) is proportionate (good balance of descriptive and analytical information)
- Detailed information and technical analysis are left for the appendix; information overload is avoided in the report
- The report provides a proper focus of truly relevant issues
- Written style and presentation is adapted for the various relevant target readers; the evaluator shows awareness of potentially different needs/interests
- Other

OVERALL ASSESSMENT OF THE REPORT

The overall assessment of the evaluation report is not a self-standing criterion. Instead it summarises key elements of, and effects resulting from the eight preceding criteria. Moreover, the overall assessment needs to balance the interests of different potential readers who may have diverging information needs:

- does the evaluation fulfil contractual conditions? (*certain internal users*);
- are the findings and conclusions reliable, and are there any specific limitations to their validity and completeness? (*most internal and external users*)
- notwithstanding intrinsic weaknesses, is the information in the report -or parts of it- a useful input for designing intervention, setting priorities, allocating resources or improving interventions (*certain internal users*)?

B. QUALITY ASSESSMENT FORM

Title of the evaluation ...

- Draft final report...
- Final report

DG/Unit

- Official(s) managing the evaluation .

Evaluator/contractor

Assessment carried out by^(*):

- Steering group
- Evaluation Function.....
- Other (please specify).....

(*) Multiple crosses possible

Date of the Quality Assessment

(1) RELEVANCE

Does the evaluation respond to information needs, in particular as expressed in the terms of references?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual** (such as deficient terms of references) and **contractual constraints** (such as lack of time, insufficient resources)

(2) APPROPRIATE DESIGN

Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual** (unexpected issues) and **contractual constraints** (such as lack of time and resources)

(3) RELIABLE DATA

Are data collected adequate for their intended use and have their reliability been ascertained?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual** (such as lack of data or access to data base) and **contractual constraints** (such as lack of time and resources)

(4) SOUND ANALYSIS

Is information systematically analysed to answer evaluation questions in a valid manner?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual** and **contractual constraints** (such as lack of resources and time)

(5) CREDIBLE FINDINGS

Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

(6) VALID CONCLUSIONS

Are conclusions non-biased and fully based on findings?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

(7) HELPFUL RECOMENDATIONS

Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

(8) EXECUTIVE SUMMARY

Does the executive summary provide a synthesis of the main findings of the report and bring out the key messages of it?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

(9) CLARITY

Does the report adequately describe the policy being evaluated and its context? Is it well structured and written in an understandable manner?

SCORING	Poor	Acceptable	Good	Very Good	Excellent
	<input type="checkbox"/>				

Arguments for scoring:

If relevant: **Contextual and contractual constraints**

**OVERALL ASSESSMENT
OF THE FINAL EVALUATION REPORT**

Is the overall quality of the report adequate, in particular:

- Does the evaluation fulfil contractual conditions?

- Are the findings and conclusions of the report reliable, and are there any specific limitations to their validity and completeness?

- Is the information in the report potentially useful for designing intervention, setting priorities, allocating resources or improving interventions?

Given the contextual and contractual constraints encountered:

- What lessons can be learned from the evaluation process?

ANNEX 9. EVALUATION WEBSITE

1. Evaluation Website: DG Enlargement Intranet

DG ENLARGEMENT INTRANET SITE - Microsoft Internet Explorer provided by European Commission

File Edit View Favorites Tools Help

Back Forward Stop Home Favorites Refresh Print Mail Stop

Address <http://www.elarg.cec.eu.int/finances/evaluation.htm>

Fast Access to ... | Home | Contact |

DG Enlargement Intranet

[Who we are] [What we do] [How we do it] [Staff Matters] [IT] [Financial Assistance]

Ex-Ante evaluations
Interim and Ad-hoc Evaluations
Ex-post Evaluations
Self-evaluations

Evaluation of Financial Assistance (E.4)

The mission of DG Enlargement evaluation function is to coordinate and implement the evaluation activities of the DG and ensure a proper follow up of the evaluations findings, results and recommendations. The evaluation function provides inputs for decision making by contributing to the design of interventions, assisting in an efficient allocation of resources, improving the quality of interventions and reporting on accountability. The basic regulatory requirements on evaluation are set out in the Financial Regulation and in the Commission Communications on Evaluation. Moreover, Enlargement evaluation guides set out the ways of planning, designing, conducting and disseminating the evaluation activities.

Head of Unit

SEGERLUND Goran - CHAR 05/204 - Tel: 92055

Documents:

- [Mission Statement](#)
- [Who does What](#)
- [Work Programmes](#)
- DG ELARG Evaluation Guides
 - [Consolidated Evaluation Guide](#)
 - [Interim Evaluation Guide Part I - Interim Evaluation Guide Part I](#)
 - [Self-Evaluation Guide](#)
 - [DG BUDG: Ex-Ante Evaluation - a Practical Guide for preparing proposals for expenditure programmes](#)

DG ENLARGEMENT INTRANET SITE - Microsoft Internet Explorer provided by European Commission

File Edit View Favorites Tools Help

Back Forward Stop Home Favorites Refresh Print Mail Stop

Address http://www.elarg.cec.eu.int/document.cfm?id_doc=22249

Fast Access to ... | Home | Contact |

DG Enlargement Intranet

[Who we are] [What we do] [How we do it] [Staff Matters] [IT] [Financial Assistance]

Documents explorer | Login in User Power/Administrator

- [-] Croatia
 - [+] Economic and Social Cohesion (+)
 - [+] Energy and Environment (+)
 - [+] Internal Market, Competition and Agriculture (+)
 - [+] Justice and Home Affairs (+)
 - [+] Social Sector (+)
- [-] Turkey
 - [+] Environment (+)
 - [+] Internal Market (+)
 - [+] Political Criteria (+)
 - [+] Transport (+)

2. Evaluation Website: DG Enlargement Internet

European Commission - Enlargement - PHARE - Interim Evaluation - Microsoft Internet Explorer provided by European Commission

File Edit View Favorites Tools Help

Address http://ec.europa.eu/enlargement/how-does-it-work/financial-assistance/phare/evaluation/interim_en.htm Go Links

European Commission > Enlargement > How does it work? > Financial assistance > Phare > Evaluation

The policy

How does it work?

- Negotiations
- Stabilisation and association
- Progress reports
- Financial assistance
- Grants / Tenders
- Technical assistance

Who does what?

Projects in focus

Turkish Cypriot community

Press corner

Candidate countries

Potential candidate countries

PHARE - Interim and Ex-Post Evaluation

Consolidated Reports

- [Consolidated Summary Report](#) [2 MB] (July 2007)
- [Consolidated Summary Report Interim Evaluation of Phare support allocated in 1999-2002 and implemented until November 2003](#) [348 KB] (March 2004)
- [Consolidated Summary Report Phare Ex post Evaluation of Country Support Implementation from 1997-1998 to 2000-2001](#) [111 KB] (May 2003)

Country Evaluation Reports

- [Phare Ex post Evaluation of National and Cross Border Cooperation Programmes - Bulgaria 1999-2001 with a Brief Review of Post-2001 Allocations](#) [523 KB]
- [Phare Ex post Evaluation of National and Cross Border Cooperation Programmes - Czech Republic](#) [545 KB]
- [Phare Ex Post Evaluation of National and Cross-Border Cooperation Programmes 1999-2001: Lithuania](#) [602 KB]
- [Phare Ex Post Evaluation of National and Cross-Border Cooperation Programmes 1999-2001: Slovenia](#) [519 KB]
- [Phare Ex post Evaluation of National and Cross Border Cooperation Programmes - Romania in 1999-2001 with a Brief Review of Post-2001 Allocations](#) [2 MB]

Thematic Evaluation Reports

- [Phare Environment Thematic Evaluation Report](#) [464 KB] (May 2007)
- [Phare Ex post Thematic Evaluation Phare Cross-Border Cooperation Programmes 1999-2003](#) [349 KB] (May 2007)
- [Support to Public Administrative and Judicial Capacity in Bulgaria and Romania](#) [1000 KB] (14 December 2006)

Right-hand sidebar:

- Twinning
- Border Regions Programme
- Evaluation of Phare
- Multi-Beneficiary Programmes
- IPA questions and answers

Bottom right sidebar:

Your Institution-Building Guide to EU Legislation TAIEX

- ▣ SIGMA
- ▣ EAR
- ▣ EIB
- ▣ EBRD
- ▣ Economic reconstruction in South East Europe
- ▣ European Fund for South East Europe
- ▣ IMF

BIBLIOGRAPHY

DG Enlargement *Phare Interim Evaluation Guide, 2004*

DG Enlargement *Self- Evaluation Guide, 2006*

DG Enlargement *Practice and Management of Interim Evaluation and Monitoring, 2004*

DG Budget *Ex ante Evaluation Guide, 2001*

DG Budget *Evaluating EU Activities, A Practical Guide for the Commission Services, 2004*

DG Budget *Communication to the Commission from Ms. Grybauskaite in agreement with the president, Responding to strategic needs: reinforcing the use of evaluation, 2007*

DG Regio *MEANS Collection. Volume 1-6 (1999), GUIDE to evaluating socio-economic development, 2007*

DG Relex, DG Dev, Europeaid, Joint Evaluation Unit, *Evaluation methodology for European Commission's external assistance, 2006*

OECD *Glossary of Key Terms in Evaluation and Result Based Management, 2002*

DG Enlargement *Results-Oriented Monitoring (ROM), Annual Reports*

DG Enlargement *IPA Programming Guide Volumes I and II, 2008*

DG Enlargement *IPA Implementation Guide, 2008*