



Directorate-General for the Budget

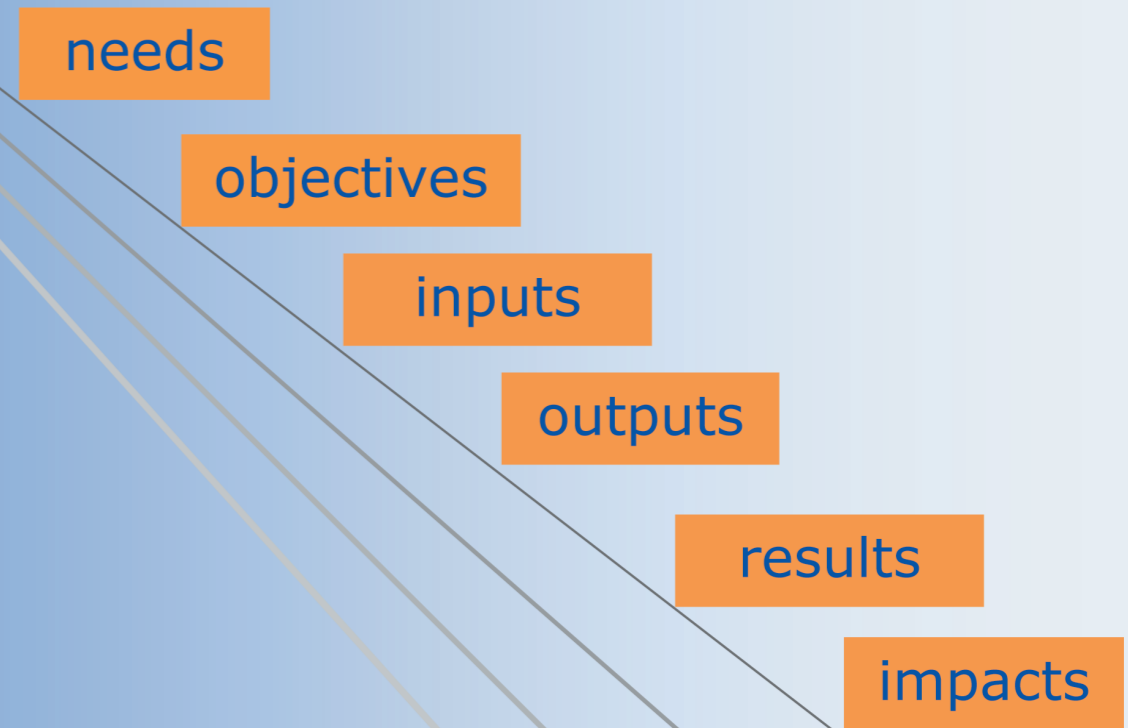
European Commission

KV-59-04-532-EN-C

Evaluating EU Activities — A practical guide for the Commission service

EVALUATING EU ACTIVITIES

A practical guide for the Commission services



Publications Office
Publications.eu.int

ISBN 92-894-7928-0



9 789289 479288

JULY 2004

SALES AND SUBSCRIPTIONS

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

How do I set about obtaining a publication?

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

How do I obtain the list of sales agents?

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

EUROPEAN COMMISSION

EVALUATING EU ACTIVITIES

A PRACTICAL GUIDE FOR THE COMMISSION SERVICES

July 2004

This guide was written by a small working group, chaired by DG BUDG. The authors wish to acknowledge helpful comments and suggestions received also from colleagues in various Commission services.

DG BUDGET – Evaluation unit

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

**Freephone number:
00 800 6 7 8 9 10 11**

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

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

Luxembourg: Office for Official Publications of the European Communities, 2004

ISBN 92-894-7928-0

© European Communities, 2004

Reproduction is authorised provided the source is acknowledged.

Printed in Belgium

PRINTED ON WHITE CHLORINE-FREE PAPER

FOREWORD

Over the past decade many important steps have been taken in the Commission to develop a regular evaluation of the Community activities. Among these are changes to the Financial Regulation, publication of Communications on Evaluation and related Standards and Good Practices, as well as the introduction to Activity Based Management.

The guide "Evaluating EU activities – a practical guide for the Commission services" responds to these developments. It provides guidance on all kinds of evaluations whether ex ante, interim or ex post evaluation and whether they concern expenditure programmes or policies.

The guide should help practitioners of evaluation in the Commission, who need to plan, coordinate, carry out or utilise evaluations. It will also assist those with a responsibility for organising the evaluation structure within the services.

I hope that this document will further consolidate the evaluation culture in the Commission and spread good practices across Commission services.

A handwritten signature in black ink, appearing to read 'Luis Romero', with a large, stylized initial 'L'.

Luis Romero

Director General of DG Budget

TABLE OF CONTENTS

INTRODUCTION.....	7
1 EVALUATION IN THE COMMISSION, ITS DEFINITION AND SCOPE.....	9
1.1 Definition and role of evaluation	9
1.1.1 Definition and purpose of evaluation.....	9
1.1.2 Characteristics and delimitation of evaluation from other information tools	10
1.1.3 Evaluation variants.....	11
1.2 Evaluation of activities.....	14
1.2.1 Activities as a centrepiece of results focused management	14
1.2.2 The scope of evaluation in the Commission	16
1.2.2.1 Delimitation of the scope	17
1.2.2.2 Evaluation under different methods of implementation.....	18
1.2.2.3 Specific cases	19
1.2.2.4 Evaluation of agencies	20
1.2.3 Evaluation in the SPP-cycle.....	21
2 PROFILE, ROLE, TASKS AND RESOURCES OF THE EVALUATION FUNCTION.....	25
2.1 Profile and role of the evaluation function in the Commission's services	25
2.1.1 Administrative set-up and role of evaluation functions	26
2.1.2 Different modes of evaluation function	27
2.2 Basic tasks of evaluation function.....	28
2.2.1 Co-ordination of evaluation activities of the Commission services....	29
2.2.2 Anticipation of the decision-making needs of Commission's services and establishment and implementation of annual plans and multi-annual evaluation programmes.....	29
2.2.3 Definition and promotion of quality standards and methods.....	31
2.2.4 Promoting the use of evaluation results in the decision-making process.....	31
2.3 Skills and resources of the evaluation function	32
3 DESIGNING AN EVALUATION	37
3.1 Mandate for an evaluation project.....	37
3.2 Evaluation questions	38
3.3 Drafting Terms of Reference.....	43
3.3.1 The purpose and justification of evaluation.....	43
3.3.2 Description of the activity to be evaluated.....	44
3.3.3 The scope of the evaluation.....	44

3.3.4	The main evaluation questions.....	44
3.3.5	The approach for data collection and analysis.....	45
3.3.6	The work plan, organisation and budget of the evaluation	46
3.3.7	Selection, exclusion and award criteria (in the case of external evaluator).....	47
3.3.8	The structure of the final report	49
4	CONDUCTING EVALUATIONS	51
4.1	Steering Groups.....	51
4.1.1	Composition of the Steering Groups.....	51
4.1.2.	Tasks of the Steering Group.....	52
4.2	Carrying out evaluations	53
4.2.1	Administrative set-up	53
4.2.2	Implementing the evaluation.....	54
4.2.3	Validation of evaluation reports.....	55
5	FINAL REPORTING, DISSEMINATION AND USE OF EVALUATION RESULTS	59
5.1	Reporting.....	59
5.2	Dissemination and disclosure of evaluation results	62
5.3	Channels for diffusing evaluation findings	63
5.4	Use of evaluation results	66
ANNEX A:	EVALUATION STANDARDS.....	67
ANNEX B:	DIFFERENT ELEMENTS AND CONCEPTS OF INTERVENTION LOGIC.....	71
ANNEX C:	DIFFERENT VARIANTS OF EVALUATION.....	74
ANNEX D:	EVALUATION OF AGENCIES	77
ANNEX E:	EVALUATION TOOLS AND TECHNIQUES.....	85
ANNEX F:	EVALUATION OF POLICIES	91
ANNEX G:	QUALITY ASSESSMENT FRAMEWORK.....	97
ANNEX H:	POSSIBLE STRUCTURE FOR EX ANTE EVALUATION REPORT.....	98
ANNEX I:	GLOSSARY.....	101

INTRODUCTION

This guide provides an overview of the Commission's rules and good practices concerning evaluation of its activities. Its aim is to help the practitioners of evaluation in the Commission, who need to plan, co-ordinate, carry-out or utilise evaluations. It should also assist those with a responsibility for organising the evaluation structure within the services.

The guide responds to a number of recent developments in the regulatory and administrative context, which have taken place since the guide "Evaluating EU expenditure programmes, ex post and intermediate evaluation" was issued in 1997. This involves in particular the Communication on evaluation¹, the Communication on Activity Based Management (ABM)², changes to the Financial Regulation³ and the Communication on Standards and Good Practices on evaluation⁴.

Moreover, a recent report⁵ on the implementation of above Communication on Evaluation recommends, in the context of improving the quality and regularity of evaluations, that DG Budget shall proceed to and co-ordinate a review of the Guide on mid-term and ex post evaluations".

This document responds to this by presenting practical solutions and good practice to help services in meeting the needs and requirements arising from the new context on evaluation. For this purpose, the structure of the document follows that of the Communication on Standards and Good Practices on evaluation, which entered into force in the 1st of July 2003.

The two first chapters provide information on the nature of and the obligation to do evaluations in the Commission as well as on how to organise the evaluation function. Together with the last chapter on using evaluation findings, they can be of particular interest to those responsible for planning, organising and co-ordinating evaluations. Chapters 3-5, which address the evaluation process (design, conduct and reporting) should be interesting for all those who have to carry-out and utilise evaluations (such as desk officers in the operational units).

Throughout this document, the standards are placed in the grey boxes in the beginning of a relevant chapter or section.

¹ SEC (2000) 1051

² SEC (2001) 1197/6&7

³ Financial Regulation (SEC (2002) 1605 and its Implementation rules (SEC (2002) 2342)

⁴ SEC (2002) 5267

⁵ The report of 10.10.2000 on the implementation of the Communication SEC(2000)1051

1 EVALUATION IN THE COMMISSION, ITS DEFINITION AND SCOPE

This chapter describes the basic objectives and characteristics as well as the scope of evaluation in the Commission. The chapter starts with a definition of evaluation and moves on to discuss the purposes, characteristics and different variants of evaluation. The second part of the chapter focuses on the scope of evaluation and how it should be understood under current regulatory requirements on evaluation.

1.1 Definition and role of evaluation

1.1.1 Definition and purpose of evaluation

According to the Communication on evaluation⁶, **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 main purposes for carrying out evaluations are as follows:**

- To contribute to the design of interventions, including providing input for setting political priorities,
- To assist in an efficient allocation of resources
- To improve the quality of the intervention
- To report on the achievements of the intervention (i.e. accountability).

In the context of the reform process launched in 2000, the Commission acknowledged the need for more results-focused management and decided, inter alia, to further develop evaluation in the Commission with the emphasis on its use as a management tool and with evaluation activities centred on activities. This process has led to a set of basic requirements on evaluation, which apply to all policy areas. These requirements are set out in the following documents:

Basic documents on evaluation:

(i) The Financial Regulation.⁷

(ii) The Implementing Rules of the Financial Regulation.⁸

(iii) The Communication on Evaluation⁹, and

(iv) The Communication on Evaluation Standards and Good Practices¹⁰.

These texts have been written at different times and circumstances, but they provide an overall and converging basis for evaluation practice in the Commission.

⁶ The Communication on Evaluation (SEC(2000) 1051): http://europa.eu.int/comm/budget/evaluation/keydocuments_en.htm

⁷ Council regulation 1605/2002, articles 27, 28 and 33

⁸ Commission regulation 2342/2002, article 21

⁹ SEC (2000) 1051

¹⁰ SEC (2002) 5267: http://europa.eu.int/comm/budget/evaluation/keydocuments_en.htm

The Financial Regulation is applicable to all Institutions. It should be noted that the Financial Regulation and its Implementing Rules are more recent than the Communication on Evaluation and normally prevail upon it¹¹, though they set an inter-institutional framework that does not prevent the Commission from defining more precise or stricter rules. The Implementing rules defines the operational, administrative or technical details enabling the implementation in practice of the principles and basic rules set out in the Financial Regulation and it must hence be read and applied in the light of the requirements of the Financial Regulation itself.

1.1.2 Characteristics and delimitation of evaluation from other information tools

Standard A.4: The role of the evaluation function shall be clearly defined and distinguished from the roles of other functions such as monitoring, control and audit.

Evaluation, monitoring and audit are three complementary but quite different exercises.

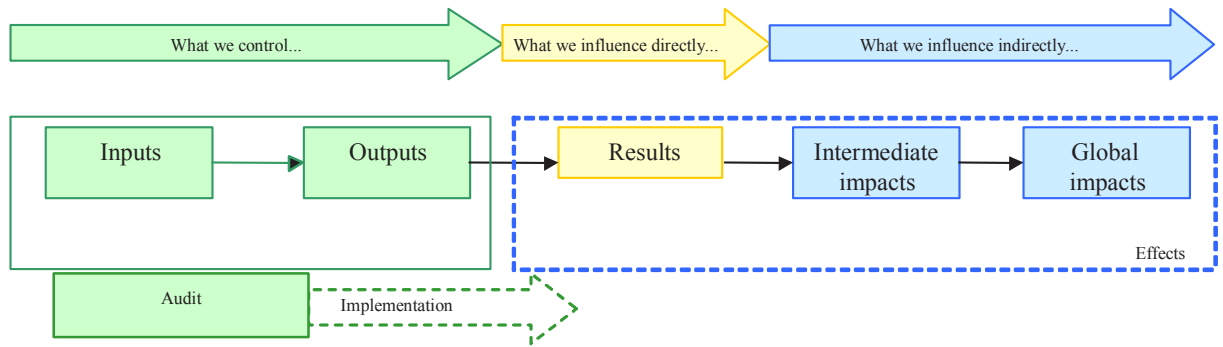
Monitoring is a continuous and systematic process carried out during the duration an intervention, which generates quantitative data on the implementation of the intervention, but not usually on its effects. The intention is to correct any deviation from the operational objectives, and thus improve the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

Audit, in the public sector covers a broad range of activities ranging from the traditional financial audit, which concentrates on inputs and outputs, to performance audit, which may encompass some features of an evaluation. A comparison of the scope of evaluation, monitoring and financial audit is presented schematically in the diagram below.

In the diagram it can be seen that different forms of evaluation cover the whole range of the intervention logic. The ex ante evaluation will, for example, address the required level of resources (input) while the interim and ex post evaluations have more focus on the results and impacts.

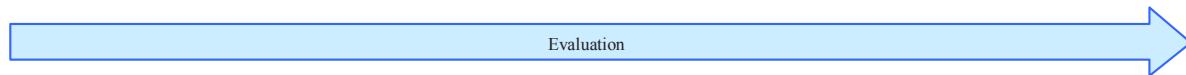
¹¹ The Implementing Rules to the Financial Regulation are for example less flexible than the Communication on evaluation concerning the frequency/timing of evaluations.



Ensuring the i) reliability and integrity of information, ii) compliance with policies, iii) safeguarding of assets, iv) the economical and efficient use of resources, and v) the accomplishment of established objectives



Assessing progress in implementation for management purposes



Judging an intervention on the basis of results and impacts at the level of addressees

Differentiating evaluation from audit

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*, *utility* and *sustainability*.¹² (For more on these concepts, see section 3.2 and Annex A).

1.1.3 Evaluation variants

Evaluation can be categorised according to different criteria, such as temporal scope, how it is conducted (internal/external) or the purpose (summative/formative).

¹² The basic documents on evaluation list some of the elements that an evaluation is expected to examine. The Standards (section 3) mentions for example effectiveness, efficiency, relevance, coherence, economy, added value and sustainability. The Financial Regulation (art. 27) brings evaluation into the context of sound financial management for which it put emphasis the concepts of effectiveness, efficiency and economy. The Implementing Rules for the Financial Regulation (art. 21(2)) imposes evaluative terms that are equivalent to effectiveness and efficiency.

Temporal variants:

As for the temporal scope there are three main variants of evaluation mirroring the programme/policy cycle. These are described briefly here. A more thorough discussion of their characteristics can be found in Annex C.

Ex ante evaluation is a process that supports the preparation of proposals for new or renewed Community actions. Its purpose is to gather information and carry out analyses which help to ensure that the delivery of policy objectives will be successful, that the instruments used are cost-effective and that reliable evaluation will be subsequently possible.

An ex ante evaluation should be seen as an analytical process that is not necessarily dissociated from the process of preparation and design of an intervention. Furthermore, different steps can be done separately. An ex ante exercise is therefore not necessarily a one-off project, with a report as the only output, but it should also be an iterative learning process. It may consist of separate phases and different pieces of analysis the results of which are reported on and summarised in the Legislative Financial Statement and in the Explanatory Memorandum (c.f. 5.2).¹³ As far as the scope of ex ante evaluation is concerned, it is specified in the Implementing rules of the Financial Regulation as follows:

The scope of ex ante evaluations:

The Implementing rules of the Financial Regulation, art. 21, specifies the scope of ex ante evaluations in the Commission by stating that:

“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, which shall identify:

- a) the need to be met in the short or long term;**
- b) the objectives to be achieved;**
- c) the results expected and the indicators needed to measure them;**
- d) the added value of Community involvement;**
- e) the risks, including fraud, linked with the proposals and the alternative options available;**
- f) the lessons learned from similar experiences in the past;**
- g) the volume of appropriations, human resources and other administrative expenditure to be allocated with due regard for the cost-effectiveness principle;**
- h) the monitoring system to be set up.”**

For more information on how to carry out above elements of ex ante evaluation, see DG Budget’s Ex ante Evaluation Guide¹⁴.

The recently launched **Impact Assessment process**¹⁵, with its special focus on policy initiatives, addresses many of the same elements as an ex ante evaluation. Both ex ante evaluation and Impact Assessment are carried out before a proposed activity is adopted. They are thus an integral part of the formulation of an intervention. Ex ante evaluation applies to all activities, whereas the Impact Assessment applies to regulatory proposals or other proposals

¹³ It is recommended to always summarise all the relevant information related to an ex ante evaluation into a specific document containing detailed information on various aspects of evaluation.

¹⁴ http://europa.eu.int/comm/budget/evaluation/pdf/ex_ante_guide_en.pdf

¹⁵ Communication on Impact Assessment COM(2002) 276 final

having an economic, social and/or environmental impact and that are presented by the Commission in its Annual Policy Strategy or its Work Programme.

The appraisal of policy initiatives (see Annex F for more information) might, however, in some cases call for a wider range of alternative options for implementation and possible effects to be identified and assessed, as well as for an examination of allocative/distributional effects.

An interim evaluation¹⁶ examines an ongoing activity whether this is a programme of limited duration or a policy, which continue for an indefinite period. In particular for the first case two specific cases of interim evaluation are commonly recognised for example in the context of the Structural Funds; namely a mid-term evaluation, which is carried out at the half-way stage of the intervention and a final evaluation, which is conducted towards the end of an intervention. These evaluations can build on the work of an earlier ex ante evaluation by assessing the continued relevance of objectives, and, in the case of expenditure programmes with a limited duration, also posing questions relating to the delivery of initial outputs and the early effects of an intervention. As far as policies with non-restricted duration (such as the Common Agricultural Policy) are concerned, the interim evaluation actually contains many elements of ex post evaluation.

An interim evaluation has also 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. In this context an interim evaluation may explicitly address some of the issues covered by ex ante evaluation.

Ex post evaluation embraces the entire intervention period, with a special interest 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. The results of a purely ex post evaluation are primarily used for accountability purposes (i.e. to provide information on what has been achieved and at what cost). However, they can also be used as input into modifications to follow-on interventions.

The Implementation Rules of the Financial Regulation provide the minimum requirements on interim and ex post evaluations. Article 21(2) states that all programmes or activities shall be the subject of an interim and/or ex post evaluations in terms of human and financial resources allocated and the results obtained in order to verify that they were consistent with the objectives set.

The Financial Regulation (in article 27) requires the Institutions to undertake ex ante and ex post evaluations, but the Implementing Rules of the Financial Regulation (in article 21) recognises that the latter will for practical reasons sometimes has to be in the form of interim evaluation. This can arise either because a decision is needed on the possible renewal of a multi-annual programme or because the activity is financed on an annual basis and a priori is not delimited in time.

¹⁶ Also referred to as intermediate and mid-term evaluation.

Internal vs. external evaluation

Evaluations may be conducted either by persons from within an organisation responsible for the intervention or by persons from outside it. The first instance is referred to as **internal evaluation** and the second as **external evaluation** (more on this, see section 4.2.1 – Administrative set-up).

Different evaluation situations are more or less well suited to internal or external evaluation. The use of external contractors is particularly suitable for evaluations with a strong summative dimension (i.e. an evaluation mainly concerned with determining the effectiveness of an intervention for the purposes of accountability or to assist in the allocation of budgetary resources) since external evaluators view an intervention with a certain degree of detachment, they may be more independent in their assessment of it. However, the possible use of external contractors is not the only factor, which affects the **independence** of the evaluation process. Other relevant factors include the evaluator's competence and integrity, his access to data and dialogue with relevant stakeholders, and also the ability to conclude freely on the basis of analysis made. It should also be noted that the independence is not an end in itself, but rather a means for achieving reliability.

On the other hand, internal evaluation carried out by the DG's evaluations function and involving operational staff is particularly useful where the formative dimension (i.e. an evaluation concerned with examining ways of improving and enhancing the implementation and management of interventions) is dominant (e.g. ex ante evaluation, and to a lesser degree intermediate evaluation).¹⁷

Sometimes it might be useful to combine both external and internal evaluation approaches within the evaluation process. For example, outsourcing the whole evaluation process to external experts may be impractical for different reasons, such as confidentiality and because direct feedback of findings into programming is important. This is especially the case in the ex ante evaluations.

However, even if the co-ordination and the main parts of an evaluation are carried out in-house, some parts of it can still be entrusted to an external expert. Examples of these are carrying out surveys or organising workshops, where the assumptions and opinions of different stakeholders are tested.

1.2 Evaluation of activities

1.2.1 *Activities as a centrepiece of results focused management*

When launching its reform in 2000, the Commission acknowledged the need for more results-focused management and decided to develop this in the conceptual framework of Activity Based Management. It regrouped all its work into a set of politically meaningful *ABB-activities* that are planned, budgeted, monitored and reported on in a new annual strategic planning and reporting cycle. This allows the systematic linking of the Commission's work with the corresponding political and managerial objectives and with the human and financial resources allocated to it. Hence, an activity becomes the centrepiece, with respect to which objectives are

¹⁷ Guide for the evaluation of activities of DG ADMIN, March 2002

defined, resources allocated and managed, and performance monitored, reported and evaluated.¹⁸

The ABB-activities (altogether 215) have been grouped into some 30 Policy Areas, closely identifiable with DGs. All Commission resources, be they financial interventions, management & support, and human resources are attributed to Activities, thus providing the full cost and its breakdown of all the Commission's Activities¹⁹.

As a response to the Activity Based Management initiative, the basic texts on evaluation have also centred evaluation on activities. The Communication on evaluation from 2000 (section 2.2.1) specifies that evaluation must be centred on *activities* and the Financial Regulation and its Implementing Rules refer to programmes or *activities* as the focus of evaluation.

On this background, the different treatment of spending and non-spending within evaluation has become increasingly irrelevant, because spending programmes and purely regulatory proposals have been merged into activities. **For a more thorough description of certain specificities concerning evaluation of non-spending activities and policies, see Annex F, where additional text on this subject is provided.**

However, since an ABM-activity corresponds to a chapter in the ABB nomenclature²⁰, it will most often embrace a significant part of an EU policy and due to the complexity of such activities it will be necessary in practice, for complex activities, to carry out the evaluation exercise at a **disaggregated level** such as actions/themes/budget lines or some other meaningful entity. However, it is necessary in such cases to ensure that the individual evaluations carried out within an activity also examine how these constituting elements **contribute to the overall objectives of this activity** (cf. the description of the activities in the Activity Statement accompanying the annual budget and the Annual Management Plan) and this information ought to be synthesised to the overall level of the Activity at a periodicity not exceeding six years as required by the Communication on evaluation from 2000.

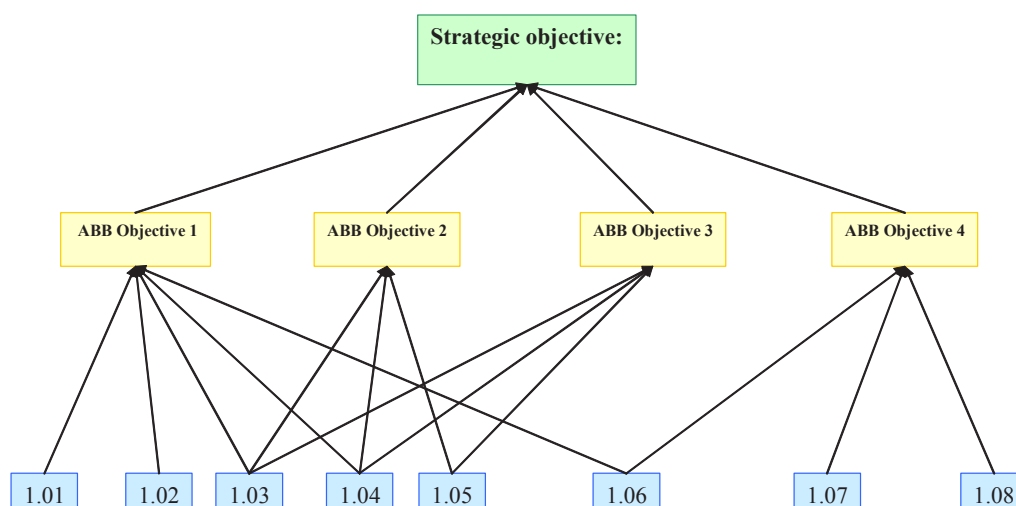
It is recommended to delimit the entity to be evaluated when planning evaluations of activities (cf. section 2.2.2) is to establish coherent groups of sub-activities for evaluation, group together those of a similar nature. Another approach is to analyse the activity tree of a DG and to cluster sub-activities according to the objectives to which they are planned to contribute, hence providing a clear link between the evaluation and reporting on the attainment of a particular objective. An example of this approach is presented schematically in the diagram below.

¹⁸ According to the ABM Glossary, an activity is defined as: “a coherent area of actions with objectives and resources, corresponding (or equivalent) to the ABB level. Each activity is related to only one policy area”.

¹⁹ The full list of activities is at: <http://www.cc.cec/home/dgserv/sg/i/spp/docs/abm/ABB.pdf>

²⁰ In this context, note that the Standards on evaluation state that they concern activities according to the ABB nomenclature.

ABB Activity 1



- Evaluation of sub-activities/actions 1.01, 1.02, 1.03.XX, 1.04.XX & 1.06.XX with respect to the attainment of ABB Objective 1.
- Evaluation of sub-activities/actions 1.03.XX, 1.04.XX & 1.05. with respect to the attainment of Objectives 2 & 3.
- Evaluation of sub-activities/actions 1.06.XX, 1.07 & 1.08 with respect to the attainment of Objective 4

Examples of sub-activities/actions are programmes (i.e. sets of actions delimited in time and resources, which may encompass several different projects, measures and processes directed towards the achievement of specific objectives; moreover, programmes have a definite time schedule and budget), pilot projects or preparatory actions. In other cases sub-activities/actions will correspond to the development and implementation of policy instruments and strategies, with or without budgetary implications.

However, it should be noted that in parallel to the centring of evaluation on activities, more broadly focused evaluations of a strategic nature are also being conducted which cut across activities and policy fields (i.e. of major legislative acts, clusters of legislation, as well as thematic priorities such as the environment, SMEs, gender, etc.).²¹

1.2.2 The scope of evaluation in the Commission

The principles enshrined in the basic documents on evaluation Commission²², require the services to carry out evaluation for activities that are directed to addressees or beneficiaries outside the European Institutions, whether or not the disbursement of funds is involved. However, these documents also foresee a certain degree of **proportionality** in the effort that should be spent on evaluation. In particular the Financial Regulation (article 27(4)) limits the obligation to evaluate to programmes and activities that entail **significant** spending and the Communication on Evaluation (section 2.2.1) states that the *scope, frequency and timing* of evaluations should be adapted to *decision-making needs* and to the *life cycle and nature* of each activity, as well as to the *resources* available.

²¹ See the Evaluation Standards for a full list

²² The Communication on Evaluation from 2000 (section 2.2.1); [the Internal Control Standard N° 23;] the Financial Regulation (article 27(4)) and its Implementing Rules (article 21).

However, these restrictions on the obligation to carry out evaluations will in general not concern the overall ABB-activities, which due to their aggregate nature normally are significant and must be evaluated. The question arising in this context is more about which level of disaggregation of the activity is necessary in terms of planning and carrying out of specific evaluation studies (cf. section 1.2.1) in order to respect the requirements outlined above. In particular, projects or other sub-activities that are not part of operational programmes, may in certain cases not warrant a separate evaluation but they may, on a case by case basis, be considered as part of evaluation studies embracing larger parts of the activity or they can also be looked at in a cursory and generic manner when the various evaluations that have been carried out over a period of time are synthesised at the level of the Activity as explained in section 1.2.1.

The decision whether to carry out an evaluation and the contents of it will hence depend on its usefulness (importance of the measures to be examined, potential use of the results) as well as on the likely added value and costs of an evaluation exercise compared to **other means** of obtaining information on the performance of the activities.

These general considerations are expanded in the two following sections.

1.2.2.1 Delimitation of the scope

The Communication on Evaluation from July 2000 specifies (section 2.2.1) that evaluation must be centred on activities and also the Financial Regulation and its Implementing Rules refer to programmes and activities as the subject of evaluation. The Communication on Evaluation furthermore explains that ‘this means that also those activities, which do not involve the use of operative appropriations should be covered by the general evaluation requirements described in this Communication. Logically, this concerns at least the activities directed to addressees or beneficiaries outside the institutions.

It is hence **necessary for the DGs to carry out evaluations** for the category of activities that concerns interventions **directed to addressees or beneficiaries outside the institutions whether or not money has been disbursed to them**. In practice this involves the evaluation of programmes and policies entailing expenditure (i.e., the classic domain of evaluation within the Commission) as well as evaluation of non-expenditure policies that ultimately are destined to affect citizens or business.

On the other hand there are situations where evaluation is not a priori an obligation, but **each DG need to consider which information feedback instruments will be the most appropriate** for assessing the performance of its activities. This concerns **activities/actions directed to the European Institutions themselves** (as opposed to external addressees or beneficiaries), such as the activities they carry out with a view to organising the administrative process of the European Institutions (decision-making, reporting, accounting, support functions, etc).

For the latter type of activities, it will be for each DG to consider, on a case by case basis, which instruments (performance monitoring, operational/ systems/performance audit, evaluation...) will be most appropriate for assessing and feeding back information on the performance of these activities or parts of activities. For this purpose it will be useful to distinguish between two different situations:

- a) **Activities (or parts thereof) that are carried out by a DG where other DGs and/or other European Institutions are addressed/benefiting.**

In this situation there will be cases where evaluation can provide added value compared to other tools and fulfil its genuine role of providing management and decision makers with a critical assessment of the activities, covering also the clients, causality and value judgements. This will in principle be linked to situations where (i) relevance and utility are important or where (ii) explaining, learning and improvement of performance are important. Another characteristic of such situations will be that the inputs and outputs of the activity give rise (beyond the full control of the DG—who provides the service) to apparent results and impacts. This case, where evaluation is useful, is illustrated by the following example.

Example - DG ADMIN as a service provider within the Commission:

DG ADMIN provides training to the rest of the Commission (output). DG ADMIN controls the quantity and quality of the training courses it provides to other DGs but it has only an indirect influence over the resulting take-up (did those colleagues having the biggest need and potential in relation to the realisation of the objectives of the training courses actually participate?) and even less control on the impact in the form of the actual use made by the trainees afterwards of the new skills learned and so forth. Normal evaluation concepts are very suitable for this case and DG ADMIN has consequently decided to carry out an evaluation.

- b) **Activities (or parts thereof) carried out by a DG and directed only to this DG itself.**

In this situation, most likely the added value of carrying out an evaluation will be limited compared to other approaches for obtaining information on the management and functioning of the activity in question. This means, in practice, that these activities/actions most often can be disregarded in respect of evaluation in their own right. However, this does not signify that evaluation should not deal with administrative and managerial activities. It is for example important that evaluation examines the extent to which implementing arrangements contribute to maximising the anticipated effects of the actions being evaluated or, to the contrary, unintentionally impedes them from materialising.

Example – predefined ABB/ABM activities:

The two horizontal activities ('Administrative support' and 'Policy strategy & co-ordination') that have been predefined for the Annual Management Plans in all Policy Areas do only concern the internal working of the DG. Moreover, chapter 1 of the ABB nomenclature ('Management and support expenditure') has no Activity Statement and no objectives and is not used as an activity in the Annual Management Plans. These activities would therefore in general not need to be evaluated in their own right.

1.2.2.2 Evaluation under different methods of implementation

The rules on evaluation apply regardless of the type of implementation method²³. However, the practical application can vary, for example where national or regional bodies are involved

²³ Insofar as they are embraced by the budgetary principles and its application as established in Article 2 and Title II of the Financial Regulation.

as in the case of shared or decentralised implementation²⁴. In this regard, some policy fields (such as the structural funds or rural development measures) have strong formal requirements on evaluation, including an array of ex ante, mid-term and ex post programme evaluations to be carried out mainly by the management authorities.

Such multi-annual programmes must be evaluated periodically in accordance with a timetable which enables the findings to be taken into account for any decision on the renewal, modification or suspension of the programme²⁵. Evaluations should hence be carried out when they are feasible, meaningful and/or formally required. In practical terms, in the case of programmes or interventions which are ongoing throughout more than one programming period, it may sometimes be possible to combine elements of ongoing, ex post and ex ante evaluation within a single exercise, provided that the scope and timing of this work is adequate. In addition, at the Commission level, the framework programme as a legal instrument should also be evaluated in accordance with a timetable, which enables the findings to be taken into account in the decision-making.

Moreover, in order to reduce the evaluation burden and maximise its utility, the evaluation needs should be mainly addressed at the level where they are more relevant. For instance, management issues should primarily be stressed in evaluations at the implementation level, while relevance and utility questions should be the subject of evaluations at the decision-making level or at the decision-making stage.

Finally, in order to allow for adequate scope and quality as well as the exchange of useful information between all involved, mechanisms should be put in place to reinforce collaboration and an adequate flow of information between all the management and decision-making levels. This might go from guaranteeing the availability of evaluation results to all actors concerned to the development homogeneous approaches to evaluation questions and methods so that commonalities are ensured.

1.2.2.3 Specific cases

It is not necessary to carry out an evaluation of each **individual project** under a programme or an activity, unless the sectoral regulation specifies otherwise. This is stated in the most recent version of the Implementing Rules to the Financial Regulation which hence formalise current practice. This means that a separate evaluation of each project is unnecessary, but the overall programme or activity embracing these projects must of course be evaluated.

In the case of **pilot projects** and **preparatory actions**, specific guidance is provided by the Communication on Evaluation of July 2000. Such actions should be subject to an ex post evaluation particularly if it is proposed to continue them in the form of a programme (hence providing an important element of the ex ante for the latter).

It should also be noted that the obligation to carry out an **ex ante** evaluation in some cases can be met in the context of an **Impact Assessment**. This will be commonplace in the case of non-expenditure activities directed to addressees and objectives outside the European Institutions. It will also be a possibility for expenditure programmes/activities in certain cases if an Impact Assessment has been carried out and it deals with the requirements of the Financial Regulation (as specified in article 21 of the Implementing Rules) and the Evaluation Standards in a satisfactory manner.

²⁴ Such as defined by art. 53 of the Financial Regulation.

²⁵ Article 21(2)(a) of the Financial Regulation.

1.2.2.4 Evaluation of agencies

As far as **Community Agencies** are concerned, the rules concerning evaluation depend on the Agency's legal form. The two most frequent cases are as follows (for more details, see Annex D, section 2):

- **Agencies covered by Article 185 of the Financial Regulation**, which have to carry out regular ex-ante and ex-post evaluations for "all programmes or activities which entail significant spending"²⁶.
- **Executive Agencies** set up in accordance with Council Regulation N° 58/2003, for which the *Commission* has to ensure external evaluations at three yearly intervals.

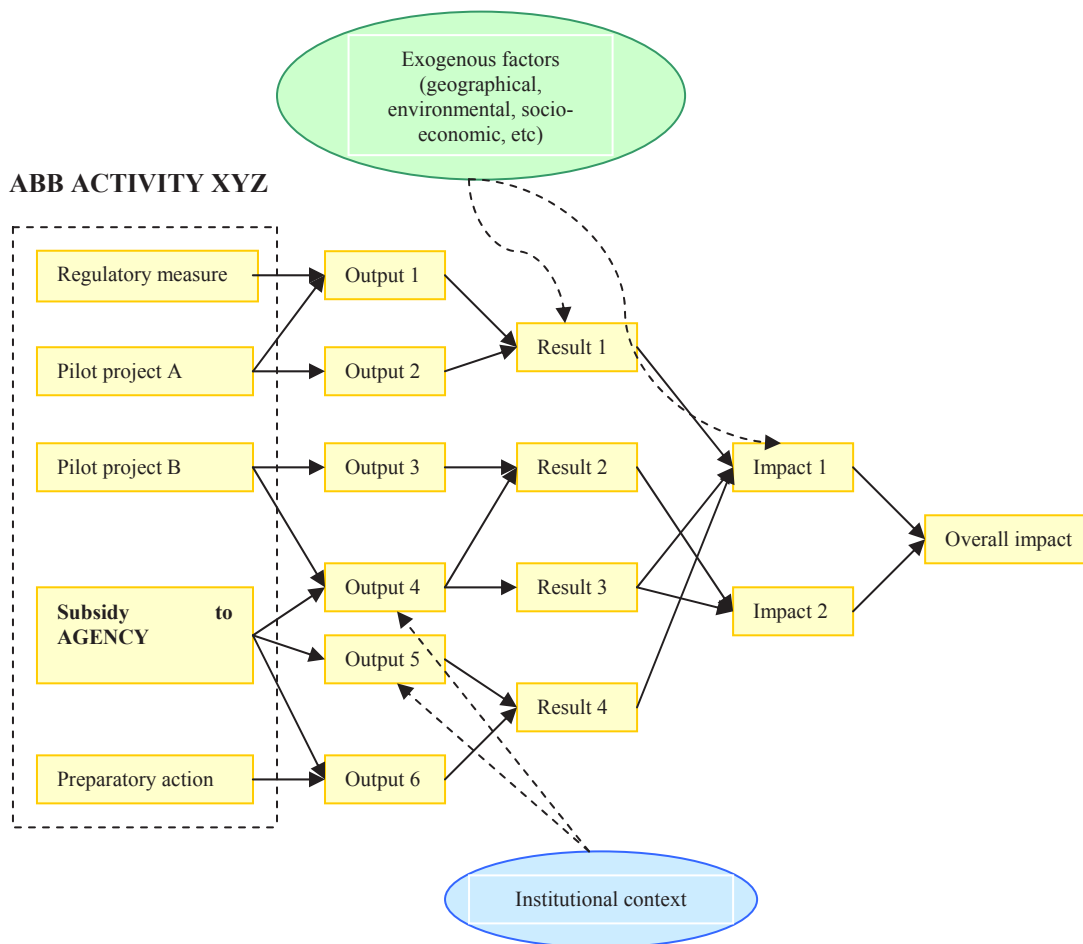
The Commission's [Internal Control Standard N°23 on evaluation](#) (like the remaining standards) does not apply *directly* to Agencies. When setting up their own control systems, Agencies have a certain room for manoeuvre in order to take account of the specific conditions under which they operate. However, the Commission's rules still serve as a *reference* (i.e. have to be taken "as a basis") for this exercise.

As far as Article 185 Agencies are concerned, some of their Constituent Acts explicitly lay down the [DG de tutelle's role](#) in the area of evaluation. In most cases, these Acts specify that it is the *Agency's* responsibility to conduct/commission evaluations, with the Commission's role being limited to the provision of information or agreement to the terms of reference proposed by the Agency.

In this context, it should also be noted that even in cases where the Constituent Acts are silent with regard to evaluation, the [DGs de tutelle are still affected by the general evaluation requirement of Article 27 of the Financial Regulation](#) (according to which evaluations have to be applied to all activities which entail significant spending): If these DGs carry out an evaluation of the ABB activity under which a particular agency is financed, evaluative information on the part of the activity relating to that agency will be required as well.

The following chart shows how the tasks performed by an Agency may interact with the other outputs produced by the DG de tutelle under the ABB activity concerned, and how the Agency's outputs may contribute to the overall impact(s) of that ABB activity.

²⁶ See Commission Regulation N° 2343/2002 of 23 December 2002.



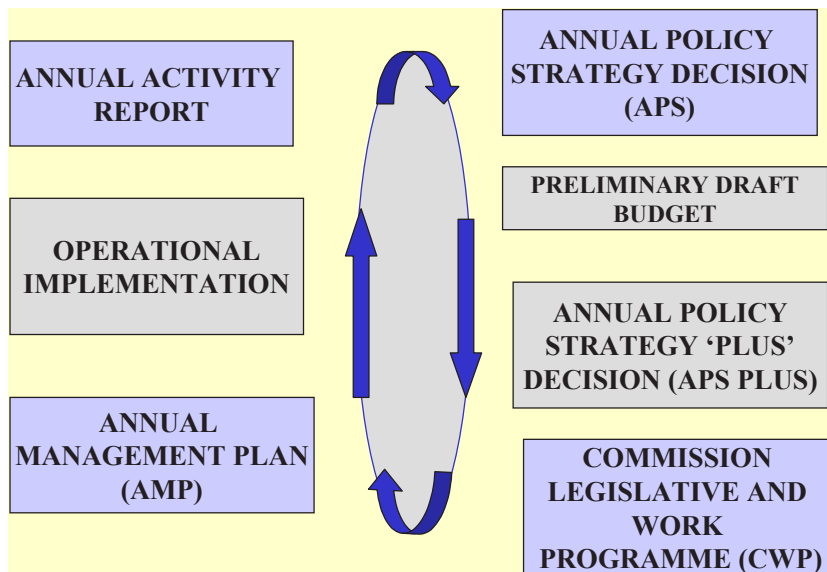
Annex D contains a recommended set of non-binding common evaluation questions that should be used as a starting point for the preparation of terms of reference for the evaluation (interim or ex-post) of Community Agencies (Section 1), as well as more detailed information on the evaluation requirements applicable to Agencies (Section 2).

1.2.3 Evaluation in the SPP-cycle

The first of the two purposes of evaluation, as presented in section 1.1.1, are closely related to the Strategic Planning and Programming Cycle (SPP cycle), which aims to link decisions on political priorities and related resources. In this cycle evaluation results are expected to provide an essential input into discussions and analyses of policy objectives, priority activities and resource requirements, as recognised in the Communication on Activity Based Management (SEC(2001) 1197/6&7)²⁷. To achieve this, evaluation must be linked carefully to the ABM activities.

²⁷ For an analysis of the experiences of how evaluation was integrated into the first two SPP cycles, see a specific document "analysis of the planning, provision and use of evaluation related information in the ABM/SPP cycle", November 2003.

The different components of SPP cycle are presented in the diagram below:



Evaluation planning and evaluation findings figure strongly at four stages of the SPP cycle

APS Decision:

The circular for the preparation of the Annual Policy Strategy indicates that:

- In the context of the multi-annual perspective of the APS, Services are expected to explain how their evaluation activities contribute to the development of their policies in the light of medium-term major changes or political developments affecting their activities.
- For each APS key initiative, Services have to present the lessons drawn from relevant evaluations that might have been carried out in this area²⁸.

Preliminary Draft Budget (PDB):

The Financial Regulation in its article 33(2) in conjunction with article 27(3) calls for evaluation results when the Preliminary Draft Budget is transmitted from the Commission to the Council and European Parliament. It states that “The Commission shall attach to the Preliminary Draft Budget...information on the achievement of all previously set objectives for the various activities as well as new objectives measured by indicators. **Evaluation results** shall be consulted and referred to as evidence of the likely merits of a proposed budget amendment”.

Inter alia, the DGs are requested to prepare an Activity Statement for each activity they manage²⁹. In the Activity Statement services are asked to provide information on **the approach, findings and recommendations of their evaluation projects**. This Activity Statement, which is annexed to the Preliminary Draft Budget (PDB), is one of the supporting documents of the resource allocation process leading to the establishment of the budget.

²⁸ This information is used, inter alia, in a synthesis of relevant evaluation findings, which DG BUDG prepares as an input into the Annual Policy Strategy Decision.

²⁹ According to a template provided in each year’s Budgetary Circular from DG Budget

Annual Management Plan (AMP):

The Annual Management Plan is an operational translation by DGs of the political priorities established in the APS. It should include the annual evaluation plan and multi-annual evaluation programme in accordance with a pre-defined template. (For more on evaluation plans, see Chapter 2 and section 2.2.2 - Anticipation of the decision-making needs).

Annual Activity Report (AAR):

In the context of the preparation of the Annual Activity Reports, the services should analyse results in terms of achievement of objectives fixed in their Annual Management Plans. Available **evaluation material** should be used, when results are being analysed.

Below is a table format presentation on how evaluation provides input for the SPP cycle in the course of a calendar year:

Year n	SPP deliverables	Evaluation planning and findings
January	AMP (year n)	Annual Evaluation plan and Multi-Annual Evaluation Programme integrated into to the Annual Management Plan of the DG.
February	APS (year n+1)	Indication of evaluations to be carried out in a multi-annual perspective. Feedback from the previous evaluations into the APS proposals at DG level. Cross-cutting input of relevant evaluation findings at the Commission level.
February		Commission's evaluation plan and programme
April	PDB (year n+1)	Evaluation input into the establishment of the Preliminary Draft Budget
April	AAR (year n-1)	Reporting on the evaluation results in the Annual Activity Report

2 PROFILE, ROLE, TASKS AND RESOURCES OF THE EVALUATION FUNCTION

The structure and content of the three sections of this chapter is based on Section A ("Profile, role, tasks and resources of the evaluation function") of the Communication on Evaluation Standards. It aims at providing guidance and practical examples on how to implement the rules and recommended good practice contained in this chapter.

2.1 Profile and role of the evaluation function in the Commission's services

This section addresses both the profile and role, which evaluation function plays in a DG or service. The profile is approached through the administrative set-up of the evaluation function in the DG and through different modes, which can be employed when the evaluation function is set up. The role, on the other hand, is explained through a short description of the rationale of evaluation in the DG.

Standards

A.1: The evaluation function shall be clearly visible in the DG's structure, the organisation chart or *Guide des Services* giving details of the unit, sector, coordinator, network members and person(s) responsible.

A.2: Evaluation shall be clearly identified in the list of each DG's operations so that the human and financial resources set aside for evaluation can be easily identified.

A.3: Each DG or Service shall define the tasks, responsibilities, organisation and procedures for running, consulting and informing the evaluation function.

A.5: Through the evaluations, the evaluation function provides support in the formulation of policies and programmes and also in management processes.

Good practices

A.1: In each DG or Service, the senior staff should take steps to provide sound support for the fundamental role played by the evaluation function in the planning of policies, programmes and all kinds of activities.

A.2: To ensure that the evaluation function contributes to the framing of policies and the setting of their objectives, it should be closely associated with the persons responsible for framing and taking policy decisions.

A.3: Where organisation is concerned, it is advisable that DGs or Services have a specific sector or unit for evaluation. If that is impossible, it is expedient to associate it with the strategic planning and programming function.

A.4: It would be advisable for the application procedures for these standards in each DG or Service to be set out in an internal document such as an evaluation agreement.

A.5: The evaluation function should encourage the departments to set in place data collection systems to assure the availability of reliable and relevant information to back up evaluation activities

2.1.1 Administrative set-up and role of evaluation functions

The principle to establish or have access to an evaluation function was already laid down in the 1996 Communication on evaluation (“Concrete steps towards best practice across the Commission”) and this principle became an obligation with the Communication on Evaluation (SEC(2000) 1051).

The term "**evaluation function**" covers the forms of organisation defined by the DGs and Services³⁰, which enable them to fulfil responsibilities for co-ordinating, framing and exploiting their evaluation activities. It is thus imperative that each DG clearly set out which form of organisation (unit, sector, network or an individual) or combination of these forms has been decided for the evaluation function of the DG and the specific role played by each of the components. The specific composition of the evaluation function of the DG can be spelt out in an operational charter for the evaluation function.

In the Commission, evaluation is a decentralised responsibility which means that the DG responsible for a given policy area is also responsible for planning, organising and managing its evaluation. Horizontal DGs (namely, DG BUDG and SG) are responsible for supporting measures. Furthermore, both horizontal DGs have a role to play when evaluation results are fed into the SPP cycle and budgetary process.

In terms of organisation chart and administrative set up, it is important that the evaluation function is clearly **visible** in the DG’s organisation chart (as mentioned in the Standard A1). For example, the evaluation function could be located under a horizontal setting³¹ within the organisation chart. In practice it can be connected to an entity responsible for policy co-ordination, for policy development and analysis, or for resources. In this context it is also advisable to foresee close links between the evaluation function and the strategic planning and programming function of the DG, (cf. Good Practice A3).

Example – linkage to strategic planning and programming on DG ENV:

The evaluation function is embedded in the Strategic Planning and Programming unit. This unit is directly attached to the Director General. This way of organising evaluation activities ensures timely and effective feedback of evaluation results into the decision-making process comprising the priority setting and the planning, implementation and monitoring of both environmental policies and expenditure programmes.

The place of the evaluation function within the organisation can also have an influence on how easily the evaluation function can contribute to the different overall goals of evaluation (see section 1.1.1). If, for example, the evaluation function needs to focus on policy formulation and development, then a set-up under a horizontal directorate dealing with policy planning or co-ordination would seem natural.

Concerning the definition of the role of the evaluation function and the suggestion in Good Practice A4 to set out the procedures in an internal document, several DGs have formalised these aspects in an operational charter for the evaluation function³².

³⁰ The term "DGs and Services" covers also the cases where an evaluation function is shared between a number of DGs and Services. In this case, an evaluation function covering two or more services should foresee specific co-ordination arrangements (for instance when the carrying out of the planned evaluations is decentralised).

³¹ Evaluation is a standard action of the horizontal ABB activity “Policy strategy and co-ordination for DG”. This horizontal ABB activity is common for all DGs.

³² Guide for the evaluation of activities of DG ADMIN:
<http://www.admin.cec.eu.int/abm/evaluation.htm#documents>).

2.1.2 Different modes of evaluation function

The Communication on Evaluation standards lists four possible modes of organisation for the evaluation function: i) unit, ii) sector, iii.) network or iv) individual(-s), but combinations are also possible.

As stated in the Good Practice A3, the preferred/ advisable modes are either a **unit** or **sector**³³.

Networks can be active within an DG, across DGs or can also serve to liaise with the Member States.

Evaluation networks within a DG may be useful as a means to involve operational units closely with the evaluation process. They can also be a practical solution, when only a very limited amount of staff is attributed to a specific evaluation function. Such networks are normally made up of members of operational as well as horizontal units of the DG.

Example – Internal evaluation network:

In DG SANCO the responsibility for selection of subject, planning, management, dissemination and the final use of evaluation results rests with the operational directorates. They are assisted in this role by a central Evaluation Section who provides information, guidance, training and a framework structure to ensure the quality of evaluations and by an **Evaluation Network** made up of representatives of each operational directorate.

The network members have the role of co-ordinating their directorate responses to evaluation issues, disseminating information from and to the Evaluation Section and providing guidance, if required, to evaluation managers.

The Evaluation Section organises bi-monthly meetings (Luxembourg and Grange participate using videoconference facilities) with a standard agenda, to discuss general evaluation issues and how to develop and improve the evaluation framework of the DG.

Whatever the arrangement mode is, the evaluation function should also be able to plan and draw evaluation findings possibly from individual evaluation reports, together at the overall level of Activity as explained in section 1.2.1, and thus contribute to the strategic decision-making within the DG.

Inter-service networks or external evaluation networks can be useful when there are strong links across the policy areas managed by more services or where activities are carried out in a devolved manner or in partnership with Member States or other external stakeholders. The latter is discussed in more detail under ‘evaluation resources’ in section 2.2.3

Example – External evaluation network:

The DGs dealing with Structural Funds have set-up external networks. Within their specific policy fields (regional policy, employment and social affairs etc.) they co-ordinate and chair the work of specific Technical groups related to the evaluation of Structural Funds. These technical groups are open to the officials from all the Member States with an aim to disseminate and share good practices. There are also external networks in the area of research, enlargement, development aid and budget.

³³ The study on the use of evaluation found that there is a strong correlation between evaluation being managed through a dedicated unit, the existence of a formal evaluation mandate and an overall evaluation plan for a DG, the elaboration of which lies with the evaluation unit.

For more on the different modes of evaluation function, including statistics of how services have organised themselves, see Annex 1 of the report below on the implementation of the 2000 Communication on Evaluation³⁴.

2.2 Basic tasks of evaluation function

This section deals with the tasks that the evaluation function has to undertake according to the Evaluation Standards.³⁵

Standard A.6:

To fulfil this role, the evaluation shall as minimum:

- a) co-ordinate the evaluation activities of the DG or Service,**
- b) help to anticipate decision-making requirements in the fields covered by the DG or Service,**
- c) establish a multi-annual evaluation programme and an annual evaluation plan for the DG or Service consistent with decision-making requirements and accordingly 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,**
- d) co-ordinate and monitor implementation of this programme and this plan at central level; where appropriate, it shall assist in decentralised implementation of this programme and this plan,**
- e) define quality standards for evaluation activities on the basis of these standards and, where appropriate, its own quality requirements, more specific and detailed or better adapted to its policy area,**
- f) promote the methodological quality and consistency of the evaluation activities of the DG or Service,**
- g) check on how the evaluation recommendations have been taken into account and,**
- h) ensure that the evaluation results can be useful in defining departmental priorities and in improving the quality of their policies, programmes and activities in such a way that these results can also be used in defining the Commission's political priorities and budgetary decisions.**

Good practices

A.6: On the basis of the evaluation results, the evaluation function should provide an independent opinion on the relevance, consistency, economy, efficiency, effectiveness, added value and sustainability of the policy, programme or activity evaluated in the light of its objectives.

A.7: In the decision-making process, the evaluation function should be consulted to take account of the results of ex ante, interim and ex post evaluations, other studies providing information useful for analysis of the problems at issue as well as past experience.

A.8: The evaluation function should systematically promote the quality of the evaluation project commissioned.

³⁴ <http://www.cc.cec/budg/networks/i/evaluation/en/notices/SEC-2000-1051.pdf>

³⁵ Furthermore a working group under the Evaluation network has identified a number of more operational tasks and activities for the evaluation function (Source: "Generic job description for evaluation staff")
<http://www.cc.cec/budg/networks/i/evaluation/en/notices/index.htm>

2.2.1 Co-ordination of evaluation activities of the Commission services

The Communication on Evaluation Standards defines an "evaluation activity" as the preparation and management of evaluation in a centralised or decentralised manner and also the utilisation of results. The co-ordination activities will in particular concern:

- Participation in and contributions to i) the Commission's Evaluation Network and inter-service working groups on thematic issues and ii) the steering committees on evaluations launched by its own service or by other services, including cross-cutting and strategic evaluations carried out at Commission level.
- Preparation of the DGs Annual as well as Multi-Annual Evaluation Plans, as well as co-ordination, monitoring and support of their implementation;
- Reporting on the evaluation activities of the DG (e.g. in the context of the Annual Evaluation Review, the Annual Activity Report, etc.) as well as promotion and monitoring of the follow-up and use of evaluation results;
- Contributions to the elaboration of strategic and policy documents.

2.2.2 Anticipation of the decision-making needs of Commission's services and establishment and implementation of annual plans and multi-annual evaluation programmes

Standards

B.1: The DGs shall plan their evaluation activities in accordance with:

- a. the Financial Regulation and its implementing provisions;
- b. the Communication on evaluation of 26 July 2000 (SEC(2001) 1051) extending the principle of evaluation to all activities in the context of activity-based management and budgeting (ABM/ABB);
- c. the rules on evaluation in the Communication on implementation of activity-based management (ABM) of 25 July 2001 (SEC(2001) 1197/6&7)
- d. the Commission's rules on impact assessment³⁶.

B.2: This planning shall take the form of a multi-annual evaluation programme and an annual evaluation plan.

B.3: The multi-annual evaluation programme shall be drawn up on the basis of the life cycle of the policies and programmes and in accordance with the nature of the activities of the DG or Service. The programme shall be approved by the Director-General -or, if the specific decision is taken, the Commission- and reflects future political priorities and operational needs.

B.4: The multi-annual evaluation programme shall spell out the division of tasks between the evaluation function and the operational departments in the centralised or decentralised implementation of the evaluation activities.

B.5: The annual evaluation plan shall be drawn up by the evaluation function and used to implement the multi-annual evaluation programme. The plan shall comprise an estimate of the human and financial resources needed to bring it into effect. The DG or Service shall adopt and supplement the annual evaluation plan in its annual management plan (AMP).

B.6: The audit and evaluation plans shall be clearly separated.

³⁶ COM(2002) 276 final of 5 June 2002

The Communication on evaluation from July 2000 indicated that a better integration of evaluation into decision-making is one of the means to consolidate and improve evaluation practice within the Commission. This implies that the evaluation function has to take into account the time span needed for formulation and implementation of an initiative to be able to meet the needs of policy makers in good time. This can only be achieved by drawing up multi-annual evaluation programmes that take account of, inter alia, the life cycle of the interventions.

Also the ABM Communication recognises that evaluation findings should provide an essential input into discussions and analyses of the policy objectives, priority activities and resources requirements. The input of evaluation material in the preparation of the APS is needed for each APS initiative; services are requested to present lessons drawn from relevant evaluations that might have been carried out in this area. In the context of the multi-annual perspective, services are expected to explain how their evaluation programme contributes to the development of their policies in the light of medium-term major changes or political developments affecting their activities.

The planning of the evaluation activities of each DG will take the form of a multi-annual evaluation programme and an annual evaluation plan³⁷. In practice, these will be integrated in each DGs Annual Management Plan (AMP) according to the format specified in each year AMP-circular³⁸.

When deciding on the content of an annual evaluation plan or a multi-annual evaluation programme, the following factors should always be considered:

- The policy objectives of the DG and the need to develop these as expressed in the APS priorities and key APS initiatives,
- Launch or renewal of legislative proposals,
- Minimum frequency set out in the implementing rules of the Financial Regulation (one evaluation for multi-annual programmes and, at least, one evaluation every six years for activities on an annual basis, in addition to ex ante evaluation from the outset of the activity),
- The need to draw together the results from the evaluations done at the sub-activity/programme level in order to find out the way in which they contribute to the activity level evaluation,
- Sectoral obligations (foreseen in the underlying legal basis),
- Requests by other EU Institutions or possible external interest,
- Results of preparatory actions, coming to term of a political mandate/deadline, need to verify problematic results at the time of the ex ante evaluation or impact assessment,
- Resources available in conformity with the Standard A-7.

DG BUDG then prepares a "Forward Evaluation Programme" which sets out planned Commission evaluations against the strategic priorities of the Commission. This exercise helps to focus future planning of evaluations.

³⁷ A good example of a recent (2003) Annual Evaluation Plan is from DG Enterprise, and available at: http://www.ent.cec.eu.int/units_directorates/a/1/evaluation.htm

³⁸ In respect of the AMP for 2003, the structure for these evaluation plans is available at http://www.cc.cec/home/dgserv/sg/i/spp/index.cfm?lang=fr&page=abm_doc

2.2.3 Definition and promotion of quality standards and methods

Each DG or service can directly apply the Evaluation Standards set out in the Communication and, whenever necessary, elaborate further specific quality standards better adapted to its own needs (c.f. Standards A.6.e and f). These specific quality standards cannot, however, contradict from the minimum Commission Evaluation Standards.

Such specific standards or guides drafted by the evaluation function of the DG can help to promote the methodological quality and consistency of the DG's evaluation activities by giving advice on the most appropriate methodologies in relation to specific evaluation objectives and on the preparation of terms of reference for evaluations.

Specific quality standards can also draw on possible international quality standards, relevant to the specific policy area. Examples of international standards are guidance provided by the OECD, World Bank etc. Examples of internal standards are provided in the box below.

Example – Quality control:

As far as the more *procedural elements of quality* are concerned DG SANCO exercises quality control using key decision points in the following manner:

- Requirement for the involvement of the Evaluation Section in specific stages in the evaluation process for example Terms of Reference, Steering Group selection etc.
- The Evaluation Section must be satisfied that appropriate actions have been taken which should ensure the best opportunity for a useful final product. If the Evaluation Section is not satisfied at any one of these key decision points they may request further work to be undertaken or alternatively remove the subject from the evaluation plan.

Example – Evaluation process:

DG AGRI has established and published a guide concerning its current evaluation practice (for a part of its evaluation activities). It covers all the steps of an evaluation starting from the decision to undertake an evaluation over the steering process as well as the reporting and dissemination until the follow up phase. This guide is used as a reference by the evaluation managers of the unit and it provides external evaluators with a clear picture of the practical arrangements they should expect in relation to for example the steering group and approval of reports. This guide is available at

http://europa.eu.int/comm/agriculture/eval/guide/cmo_en.pdf.

Where relevant, the evaluation function should organise training to foster the promotion and spread of methodological tools adapted to the specific needs of its DG.

Moreover, although the MEANS collection was established for the purpose of evaluation in the structural funds, it offers a good example of guidance also applicable to other sectors.

2.2.4 Promoting the use of evaluation results in the decision-making process

The evaluation function should also take steps to make sure that evaluation results are used in decision-making. This will be addressed in chapter 5.

2.3 Skills and resources of the evaluation function

This section addresses both the financial and human resources needed to undertake the tasks as described above. The aim is to give some benchmarks for the services, when deciding on resources. For this purpose, the text gives some basic requirements as well statistics and examples from current practice.

Standards

A.7: The human and financial resources available for evaluation work shall be clearly identified and planned. These resources shall be sufficient for implementing the multi-annual evaluation programme and annual evaluation plan of the DG or Service.

A.8: In relation to the evaluation plans approved, the resources earmarked for the evaluation function shall be adequate and appropriate in terms of funds, staff and skills so as to ensure that they can fulfil their responsibilities effectively

Good practice

A.9: The financial resources set aside for evaluation by each DG or Service can vary considerably, both in absolute volume and as a percentage of the total budget, depending on the nature of the activities performed and the budgets administered. However, in the case of expenditure programmes, an indicative 0.5% of the budget could serve as a basis for calculating the amount to be set aside for evaluation.

A.10: The evaluation function should have at its disposal financial resources to carry out work on methodological questions.

Evaluation skills

Staff within an evaluation function require a variety of skills to enable them to plan, prepare and manage evaluations, or to provide advice and support to staff in operational units when the latter are directly responsible for an evaluation project. They may also be called upon on occasions to conduct evaluations themselves internally or to support the operational services to do this.

In order to assess and reply to the evaluation needs of a DG, evaluation managers should have a good understanding of policy analysis and the theory and practices of evaluation as well as the legal framework for evaluation, how it is to be applied and the scope and limits of evaluation. They should be capable of analysing activities, policies and objectives as well as related information needs of a DG. They should be able to network effectively within their service and ensure that colleagues are aware of and take note of evaluation results. Finally, they should also know how to plan and manage the reporting, diffusion and follow-up of an evaluation in order to support the use of its results.

An externally contracted evaluation has to be prepared and managed in the same way as any project, with the establishment of a clear project specification in terms of the services required and the conditions in which they are to be delivered. This requires that evaluation managers have good knowledge of the evaluation process (i.e. the phases of an evaluation method, what they entail and how long they take to implement) the various constraints that weigh upon it (information availability, time and costs), the characteristics of the deliverables to be produced by the contractor, how to establish the focus of an evaluation (i.e. evaluation questions) and, last but not least, the Commission's procurement processes.

An evaluation manager also needs to be able to liaise effectively with concerned units/directorates/DGs before, during and after an evaluation. This involves a variety of skills including being able to communicate and translate evaluation concepts for operational staff and management, transpose the information needs of services into analytical form (i.e. evaluation questions), mediate between the evaluator and services (i.e. ensuring that all sides understand and fulfil their respective roles) and transpose the evaluation results into useful and usable information for the services.

Finally, evaluation managers should have a sound understanding of key evaluation concepts and evaluation design (i.e. the use of appropriate tools and techniques in a coherent way within an evaluation project) in order to be able to establish appropriate methodological guidelines for terms of reference, assess the technical quality of evaluators' offers, interact meaningfully with the evaluator over methodological choices made during the execution of the evaluation, and assess the quality of deliverables and the evaluation process from a methodological perspective. Furthermore, in the case of internally conducted evaluations, the internal evaluator should have a good level of knowledge of how to assemble and implement evaluation tools and techniques, including their strengths and weaknesses in respect to specific evaluation situations.

Evaluation resources

While it is not possible to provide precise indications of the level of **human resources** required to plan, prepare and manage evaluations within a DG, there is a clear relationship between the mode of organisation of evaluation, of which there are three main types within the Commission, and staffing needs. Where evaluation functions are themselves directly responsible for planning, preparing and managing evaluations (i.e. a centralised evaluation function), staffing levels need to be higher. The study on the use of evaluation in the Commission services indicates that five or more full-time, trained evaluation managers is usual in this case³⁹. On the other hand, when responsibility for evaluation largely resides with operational units and the evaluation function has an advice (methodology, procedures, etc.) and co-ordination type role (i.e. decentralised evaluation function), staffing levels will be lower (up to four full-time, trained evaluation managers is usually the case according to the above study).

Some aspects of the responsibility for evaluation may also be devolved to Member States or authorities in third countries, a mode of organisation that can (and does) coexist with the previous configurations. This situation requires that additional evaluation staff are available to co-ordinate and support evaluation activities undertaken by external bodies, through the development of guidelines, participation in steering groups, provision of help to line services, etc. as well as to undertake the production of evaluation syntheses of the considerable quantities of information generated by devolved evaluation exercises. This also requires that country desk staff have a good level of awareness of evaluation issues in DGs where devolved evaluation responsibility exists.

An general overview of the current situation in the Commissions concerning human resources allocated to evaluation functions is provided in the report⁴⁰ on the implementation of the Communication SEC(2000)1051. There were (in 2002) 17 DGs with less than two staff in the evaluation function, 3 between 2 and 5 staff and 4 with more than 5 staff members. However, the above figures do not include those working in the operational units who are involved in evaluation activities in addition to the staff in the evaluation function.

³⁹ The study on the use of evaluation in the Commission services was carried out by external consultants in 2002: http://europa.eu.int/comm/budget/evaluation/studies/study_en.htm

⁴⁰ Annex 1 of the report of 10.10.2000 on the implementation of the Communication SEC(2000)1051 (currently only available in Intracomm at: <http://www.cc.cec/budg/networks/i/evaluation/en/notices/SEC-2000-1051.pdf>)

As far as **financial resources** are concerned, Good Practice A-9 stresses that cost of an evaluation project can vary a great deal although it gives an indicative amount of 0.5 % of the budget allocated to an expenditure programme. The relevance of this indicator would depend on the nature of the intervention. It may be useful to spend high proportion on pilot projects and preparatory actions while for large expenditure activities (such as Common Agricultural Policy), the figure would be lower.

Good Practice A-10 notes that the evaluation function should have at its disposal financial resources to carry out work on methodological questions.

Example – Methodological development

In DG JAI, the evaluation function has been given the mission to develop an analytical framework for the evaluation of policies and legislation. This work is of horizontal nature to the DG, and requires knowledge and experience in evaluation. To support this task, a specific budget has been allocated in 2003 to a preparatory, desk-based study which is now completed. The next stage will imply testing the analytical framework in two policy areas, still under the responsibility of the evaluation function. If successful, the framework for policy evaluation will be endorsed and applied across the different areas according to the standard evaluation requirements and share of responsibilities with the operational units.

When estimating the cost of an evaluation project, several variables may have an influence on it, such as: size and duration of the intervention, scope and complexity of the intervention, geographical diversity, size and nature of the stakeholders/target population/beneficiaries, quality of monitoring systems in place and methods foreseen.

The MEANS guide (volume I, page 80 in the English version) contains the following table indicating the estimated cost of some evaluation techniques:

Technique	Cost (man/month)	Duration
Questionnaire surveys	Medium	XXX
Surveys based on individual interviews	Medium	XX
Survey based on focus groups	Low	XX
Surveys based on case studies	Medium	XXX
Statistical analysis	Low	XX
Expert Panels	Medium	XX-XXXX
Use of econometric model	Low	XX
Development of an econometric model	High	XXXXXX
X = approximately 1 month		

Each of the Commission's Annual Evaluation Reviews provide an overview of the average cost of an evaluation project as well as the variation between years and between headings of the Financial Perspectives.⁴¹ :

For example, the Annual Evaluation Review 2002 shows that the average cost of an evaluation project in several years has been in the range of 215 000 to 230 000 €.

The study on the use of evaluation in the Commission services (cf. above) also provides a typology of the DGs evaluation arrangements. It combines several of the aspects presented above. This typology (cf. section 6.16 of the report) recognises three categories of DGs: those having respectively a higher, a medium and a lower level of evaluation institutionalisation. The DGs having a higher level of evaluation institutionalisation are characterised by i) important evaluation experience, ii) sizeable, dedicated and trained evaluation function, iii) established quality control and transparent procedures and iv) full or partial feed back mechanisms in place.

⁴¹ http://europa.eu.int/comm/budget/evaluation/pdf/review2001_en.pdf

3 DESIGNING AN EVALUATION

This Chapter deals with the different elements related to designing an evaluation. It starts with a discussion on the mandate for an evaluation project, and then moves on into the more detailed aspects of the preparatory phase of an evaluation involving how to devise evaluation questions and draft of the Terms of Reference.

Standards:

C.1: Each evaluation function shall be devised as a project comprising three separate stages: design, implementation and exploitation. The project manager shall establish who is to take part in these three stages in accordance, where appropriate, with the internal rules and procedures applicable.

C.2: The points of importance to the different departments concerned shall be examined when the questions for the evaluations are established.

C.3: When designing any (internal or external) evaluation project, the purpose of the evaluation must be clearly and accurately defined.

C.4: In the event of external evaluation, the evaluation project must also include the contractual, financial and administrative clauses and quality criteria. Specifications shall formalise the above points.

3.1 Mandate for an evaluation project

As an intermediate step between the specific commitments to carry out an evaluation (e.g. sectoral regulations) and the detailed Terms of Reference, it is sometimes appropriate to establish an evaluation mandate. This mandate is a brief and overall description of the evaluation to be carried out. This document should specify what is going to be evaluated, what the motives are for doing an evaluation, how the work will be organised, and what the expected use of evaluation is. It can take the form of a specific evaluation mandate, but also other documents can lay down requirements for evaluation and provide guidance on how to do it.

The drawing up of an evaluation mandate is the first step of a well-prepared evaluation and enhances the credibility of the whole evaluation exercise. A key advantage is that it commits the internal Commission stakeholders to the evaluation process and its results, thus ensuring support and reducing the risk that they will not endorse its conclusions and recommendations.

An evaluation mandate is normally signed by the Director General. It is drawn up by the evaluation function in collaboration with operational/policy unit concerned, involving the primary users in the process of developing a formal and explicit framework for the evaluation.

Below is a list of items to consider when drafting the mandate for an evaluation:

- **Context of the evaluation project.** What are the decision-making needs of the service(s) concerned? Are important reforms or smaller modifications foreseen in the near future, for which evaluation evidence is needed? Who are the stakeholders of the evaluation?
- **Motives and objectives of the evaluation project.** A short description of possible compulsory requirements (such as the Council Regulation (1260/1999/EC) laying down the principles of the Structural Funds) or specific decision-making needs should be described as a motive for the evaluation project.

As far as the objectives of the evaluation project are concerned, they can be briefly listed in the mandate, and further developed in the Terms of Reference.

- **Calendar of the project.** A short description listing the main milestones is sufficient, because the calendar will be further specified in the Terms of Reference.
- **Management of the project.** This point should identify the following aspects:
 - Who drafts the Terms of Reference and chooses of the evaluator,
 - Who has the responsibility of the evaluation project and what is the role of the stakeholders in this respect,
 - Data collection and necessary arrangements within the service to carry it out,
 - Decision-making process with regard to the evaluation process,
 - Mechanisms for the approval of the quality of the submitted reports,
 - Arrangements for the dissemination of evaluation results,
 - Arrangements for the implementation of recommendations.

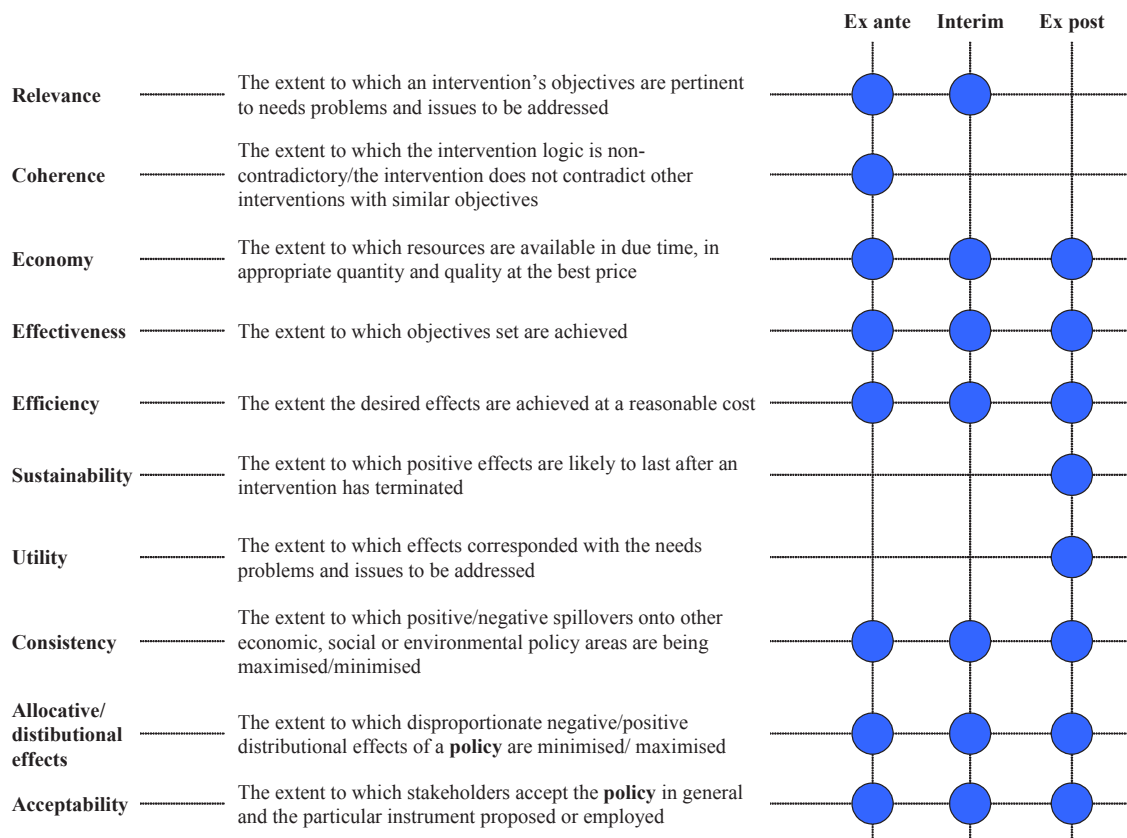
3.2 Evaluation questions

Standards

C.2: The points of importance to the different departments concerned shall be examined when the questions for the evaluations are established.

C..3.d: When designing any (internal or external) evaluation project, the purpose of the evaluation must be clearly and accurately defined. The evaluation project shall comprise the key questions.

The elaboration of evaluation questions is a part of the **structuring phase** of an evaluation (cf. section 4.2.2) in which the effects to be evaluated are chosen and clarified. Evaluations should provide essential information of an evaluative nature for their users. This is normally ensured through the drawing up of a series of questions aimed at directing the work of the evaluator. This will be the more necessary the more the intervention to be evaluated is complex in nature (i.e. the intervention is implemented through multiple instruments and/or has multiple objectives). The information a particular evaluation generates centres on a number of **key (generic) issues** embodied in the evaluation questions. The issues that an evaluation should address depend to some extent on its timing. The diagram below indicates and explains the evaluation issues (relevance, efficiency, effectiveness etc.) that are most relevant to different stages of the evaluation cycle.



This diagram gives general guidance of how to orient an evaluation. However, a good evaluation needs to deal with these general evaluation issues within the specific context of the intervention being evaluated. This means that [specific evaluation questions](#) need to be elaborated. In order to do this, it is important first to understand the intervention logic. The latter should in principle be available from the ex ante evaluation. However, in practice it may often be necessary, as a preliminary step to better identify all the objectives (general, specific, operational) and the actual outputs, expected effects (results, impacts) as well as the principal relationships between the outputs and effects. Depending on the complexity of the case it may be necessary to systematise (make typologies) of the outputs, effects and addressees or beneficiaries.

It should be noted that this preliminary structuring work need not always examine the intervention logic and causality in full depth; only to the extent needed for drafting the evaluation questions. The contractor can in later stages of the inception phase go into more depth with this work in order to provide detailed typologies, to verify the evaluation questions and to propose the criteria and indicators for answering them.

On the basis of this preliminary structuring work specific evaluation questions can be formulated and the most important ones selected for the Terms of Reference. Certain tools can be helpful in this phase in order to list questions, clarify them, grade them and select those for which the evaluation must provide answers. Some of these tools are presented in the diagram in section 4.2.2 and in ANNEX E.

The specific evaluation questions should normally be analytic, and not only descriptive. The questions would often combine different elements of the intervention logic as well as a preliminary judgement criterion. This enables verification, or maybe even a quantification of key elements of the intervention logic. This is illustrated in the box below which reproduces an evaluation question from rural development.

Example of evaluation question embracing part of the intervention logic:

To what extent have the supported investments increased the competitiveness of agricultural products through improved and rationalised processing and marketing of agricultural products?

This example illustrates a general scheme for the construction of this type of questions, i.e., have certain inputs/outputs improved⁴² the results/impacts concerning certain beneficiaries.

In addition, the terms of reference would also often include more summative questions in order to/assess the overall value of the intervention.

The following example from DG AGRI illustrates various elements of the structuring phase. It shows one out of the nine the evaluation questions, which were included in the terms of reference following the internal, preliminary structuring work. This example also embraces the second step of the structuring phase (identification of observation tools) because it also presents the indicators, which in combination with the judgement criteria were used for answering the evaluation questions in this specific evaluation (on forestation of farm land). The criteria and indicators are taken from the inception report where the contractor indicated in detail how he intended to deal with each evaluation question.

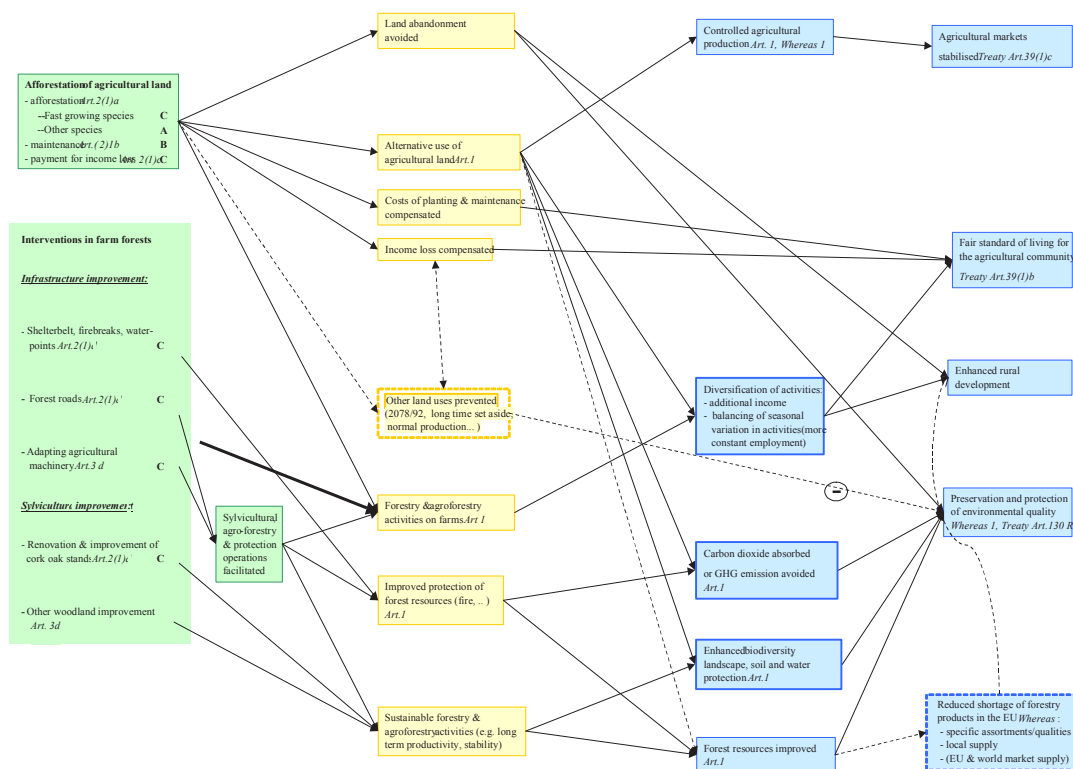
⁴² 'Improved' is vague or provisional criterion, which during the inception period should be specified in more detail

Example: Structuring phase and evaluation questions

This example from DG Agriculture⁴³ illustrates the structuring phase of an evaluation of a Regulation, which inter alia supports the afforestation of agricultural land. The example presents the preliminary intervention logic and one out of the nine the evaluation questions, which were included in the terms of reference following this analysis. This particular question corresponds to the shaded area of the diagram of the intervention logic. The example also provides the criteria and indicators, which resulted from the work in the inception phase.

EXPECTED INTERVENTION LOGIC:

Aid Scheme for Forestry Measures in Agriculture (Regulation 1080/92)



QUESTION	CRITERION	INDICATOR
To what extent has the afforestation of agricultural land contributed to reducing agricultural surpluses	Reduction of agricultural surplus which can be attributed to the Regulation	Area retracted from agricultural production thanks to afforestation through the Regulation (specified to the different types of previous agricultural land use) {hectares}
		Agricultural surplus production retracted thanks to afforestation by type of previous agricultural land use {tons}
	Compensation of the afforestation costs by a reduction in the Community costs linked to storage of agricultural surpluses	Comparison of afforestation costs with the costs of storage of surpluses (specified to type of production resulting from the previous agricultural land use) {€}

⁴³ The report is available at http://europa.eu.int/comm/agriculture/eval/reports/forest/index_en.htm

This type of structuring work can also be a useful preparation where a DG need to produce guidance on for example indicators for devolved evaluations to be carried out at country or regional level.

The performance of an intervention measured by these generic evaluation issues can also be influenced by the [implementing arrangements](#) for the intervention. These arrangements can hence be among the factors explaining, for example, the effectiveness, efficiency and sustainability of an intervention.

Example – Implementing arrangements affecting performance

The evaluation should, where appropriate, address whether the implementing arrangements are adequate for the purpose of the intervention.

However, from an evaluation perspective, the implementation arrangements should not be assessed in isolation; instead the evaluation should examine whether these arrangements would be or have been conducive to (or have hindered) achieving the intended outputs, results and impacts. The aim of the assessment would be to have a better informed choice between management options. This question is hence especially important for ex ante or interim evaluations that serve as an input to decision making on the overall management method or at least its practical modalities for a new programme, a successor programme or for the remaining programme period.

The terms of reference may therefore include questions like:

- To what extent do the organisational set-up/ management systems and processes contribute to the effectiveness and efficiency of operations? *[to be adapted to the specific case]*

As background information it should be noted that the **overall methods of implementation** available to the Commission are set out in the Financial Regulation: centralised (directly/indirectly), shared, decentralised, and joint management. In addition a large number of **implementing modalities** (projects selection, payment arrangements...) have to be organised. The key factors influencing the choice and appropriateness of a management method as well as the trade-offs to be made involve⁴⁴:

- Scale & scope of the intervention
budget, timescale, speed of implementation (innovative, market orientation, / location (EU, regional...))
- Beneficiary type
NGOs, universities, SME, large firms...
- Policy Area and type of objectives
economic, social...
- Contextual factors.

It should also be born in mind that the choice of management method and of the practical implementing modalities influences the costs incurred by the Commission and other implementing bodies, right down to the addressees/beneficiaries. Annex F explains the types of cost that might occur in addition to the budgetary costs (operational, human, administrative) at the different levels of implementation.

⁴⁴ The strategic evaluation of management methods.

In addition to the analytic questions described above an evaluation in its terms of reference will most often also include descriptive questions requiring an overall description of the intervention being evaluated, including its various inputs, outputs, addressees. A proper identification of the inputs and outputs is also a prerequisite for adequately dealing with the generic evaluation issues, in particular the efficiency of an intervention.

3.3 Drafting Terms of Reference

The Terms of Reference (TOR) is the key document in the evaluation process, especially when the evaluation is being contracted to an external evaluator, presenting the origin, scope and aim of the evaluation project, as well as the distribution of roles. The TOR should describe the content of the activity to be evaluated and specify the information, which is already available and will be made available to the evaluator. Furthermore, it should state precisely a limited number of questions, which the evaluation has to answer. It will set out the steps in the evaluation process also provide an outline of the evaluation method to be used and set the specific steps the evaluation process must follow and it may also provide an outline of the evaluation method to be used.

Below is a list of the main elements, which the TORs should include:

- The purpose, objectives and justification for evaluation (including the legal base),
- A description of the activity to be evaluated,
- The scope of the evaluation,
- The main evaluation questions,
- The overall approach for data collection and analysis,
- The framework delimiting the work plan, organisation and budget of the process,
- In the case of external evaluator, clear selection criteria,
- The structure of the final report, and if possible also of the progress reports,
- The expected use, and users of the evaluation.

The TOR provides all the key specifications for the work to be carried out including the basis on which the evaluation method is designed. This initial orientation of the work may be specified in more detail in the course of the evaluation in the light of new information available after the evaluator has been selected (the result of the structuring phase – see section 4.2.2). The various elements of the TOR figuring in the above checklist are further developed below.

3.3.1 The purpose and justification of evaluation

The Terms of Reference (TORs) usually describes the legal requirement for carrying out the evaluation (e.g. referring to the specific regulation, directive or decision requiring an evaluation within certain time frame or referring to the general regulatory requirements on evaluation, as listed in the section 1.1.1).

In addition, this part of the TORs should also explain the purpose of the evaluation, and thus define the general objectives and intended use of the evaluation project.

Example – Purpose and intended use of an evaluation

DG AIDCO Country evaluation on Ukraine states that “the objective of this evaluation is to provide the Commission with key findings and lessons of experience from the Commission’s past and current assistance to Ukraine. Taking account of Ukraine’s political, economic, social and environmental framework, it will present those findings and lessons, along with a set of detailed recommendations, in a report designed primarily to provide the Commission’s policy-makers and managers with a valuable aid both for the implementation of the current strategy and indicative programme and for further programming.”

3.3.2 Description of the activity to be evaluated

TORs should also provide basic information on the exact content of the activity (policy, programme etc.) to be evaluated. Thus, the following items could be covered:

- Problems, needs and issues the intervention is intended to address,
- Target population of the activity,
- Objectives of the intervention,
- Delivery systems (actions, budget lines etc.) currently in place,
- Outputs to be delivered (types of projects financed, services or products delivered, etc.)
- Expected results and impacts.

3.3.3 The scope of the evaluation

Depending on the nature of the activity, the scope can comprise different elements, but at least the legal, temporal and geographical aspects of the scope should be considered:

- Will the evaluation cover the entire intervention? If not, then TORs should clearly indicate which part is to be excluded (part of the budget, periods of time, selected parts of the target population, portion of geographical areas subject to the activity etc.),
- Which other interventions should be addressed in the evaluation? This can take various levels, such as:
 - other interventions within the same activity,
 - other, related activities (and their underlying interventions) of the same policy area,
 - other, related activities (and their underlying interventions) of different policy areas,
 - other, related interventions of the Member States, and
 - other developments in the context.

3.3.4 The main evaluation questions

The advantage of having a set of clear evaluation questions (c.f. section 3.2) in the Terms of Reference is that the evaluator is able to propose a method that is tailored to suit the information needs of users.

3.3.5 *The approach for data collection and analysis*

The requested 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:

- providing a focus and structure to the evaluation project
- collecting data in the field
- analysing data collected in the field and from other sources
- providing judgements

When addressing the question of the methodology in the TOR, it should be kept in mind that there are many different evaluation methods possible (i.e. combinations of evaluations tools, techniques and various approaches – see section 4.2.2). Hence, in the case of an external evaluation, it can be in the interest of the commissioner of the evaluation to provide only a general outline, and thus invite the potential evaluator to propose an appropriate solution, depending on his/her experience and expertise. Nevertheless, it can be useful, for example, to ask the contractor to develop typologies of, e.g., outputs or addressees, to specify minimum requirements concerning the representativity of the data to be collected or analysed, and to state how the analysis of causality will be approached.

Other factors also influence the design of the evaluation, notably the time and resources available for the evaluation project. Bigger budgets and longer time horizons enable the commissioner of the evaluation to ask for more varied and in-depth data to be collected and for more sophisticated analyses to be made. The possibilities and restrictions that follow from the size of the budget can be summarised as follows:

Short time span and small budget: - desk studies, interviews, focus groups.

Medium time span and budget: - case studies, surveys, expert panels.

Long time span and considerable budget: - econometric models, cost-benefit and cost-effectiveness analyses.

Finally, other aspects of an evaluation have important repercussions on the choice of the evaluation method, such as:

Does the evaluation concern a programme or an activity addressing clearly identified beneficiaries or is it a policy that addresses more or less indirectly a broad and potentially difficult-to-contact public?

Will the evaluation focus just on the direct, shorter-term effects on addressees (e.g. as in the case of a mid-term evaluation) or will more indirect effects, longer-term effects and unintended effects be addressed?

Whether or not issues such as the displacement and substitution effects are important and need to be addressed?

For a more detailed presentation of evaluation methodology and different evaluation tools and techniques, see section 4.2.2.

3.3.6 The work plan, organisation and budget of the evaluation

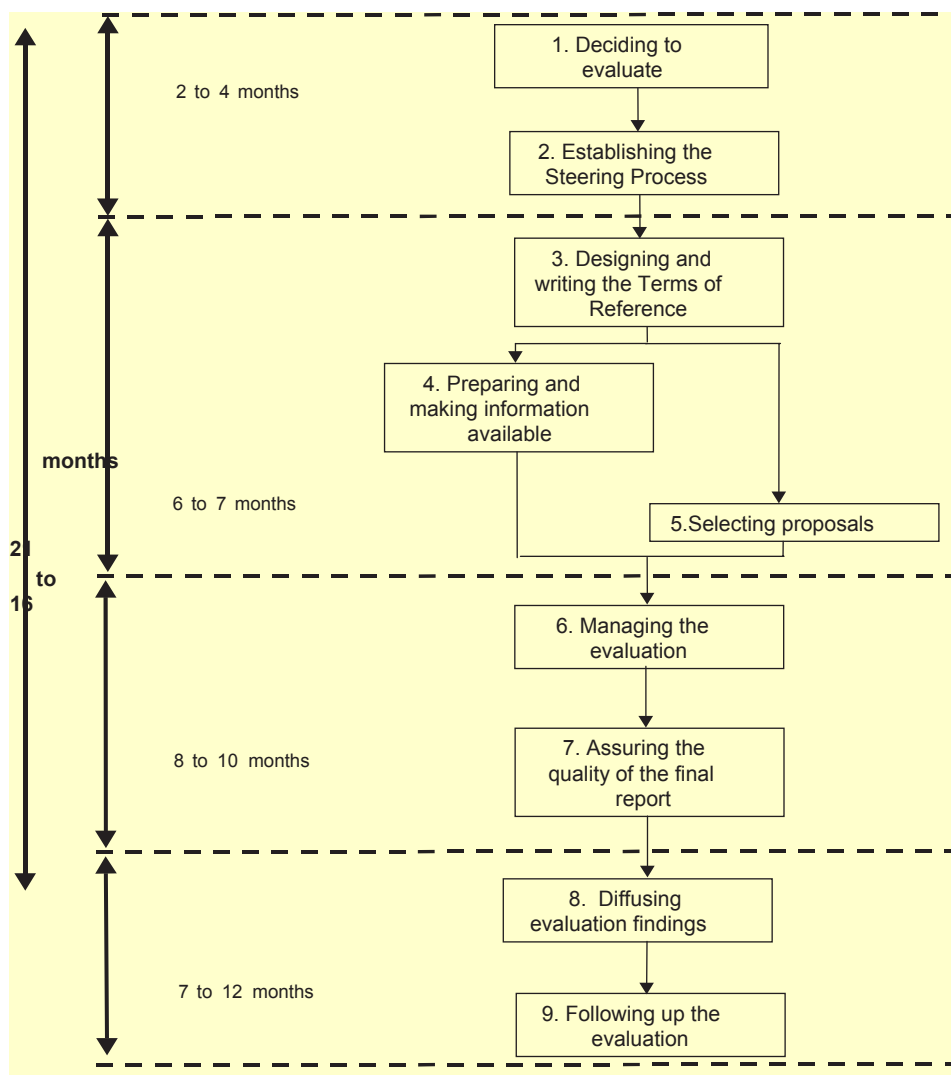
The work plan is always linked to the decision-making needs (i.e. the date by which the information from the evaluation is required by the services), and thus should never be drafted without this constraint in mind. When the link to the decision-making process has been established it is essential to establish clear deadlines for different phases of the evaluation process, the meetings with the steering group and the deliverables to be supplied by the evaluator as the work progresses. Moreover, guidance on existing data sources and relevant contacts for the evaluator should also be provided.

The organisation of an evaluation involves specifying the different roles of actors and what is expected from each of them. This point could also cover reporting responsibilities and the procedures to be followed in the reporting.

As discussed above, the financial resources available for the evaluation project have important repercussions in terms of its scope, the method used and questions that can be addressed in the course of the project. In particular, the number of questions that the evaluation can address and the level of detail in which they are addressed are in part a function of the evaluation budget. Also, procedural aspects such as number of meetings foreseen can have an important influence on the budget needed for the evaluation project. For more information of the budget of the evaluation project, see section 2.3 (resources), where information is given on the possible costs of employing various evaluation techniques.

As an example of different steps and related estimated duration of an evaluation process, consider the following work plan from DG AGRI.

Example: evaluation process duration



3.3.7 Selection, exclusion and award criteria (in the case of external evaluator)

Evaluation work is considered a service contract of an intellectual nature listed in Annex IA to Directive 92/50/EEC and governed by the procurement rules set out in the Financial Regulation and its Implementing rules⁴⁵.

There are three types of procurement procedures: open, restricted and negotiated. The procurement rules lay down which types of procurement procedures are possible depending on the nature and value of the contract. The procedure must fully comply with all the provisions established for the contracting modality chosen. The non-observance of these provisions will invalidate the call for tender.

Whatever the contracting modality, the invitation to tender (see art. 130 of the Implementing Rules of the Financial Regulation) shall have attached to it, i.a., the terms of reference (specifications). The TOR (if this is not specified in the contract notice) shall specify the following criteria:

⁴⁵ See in particular Title V "Procurement" of the Commission Regulation 2342/2002 of 23.12.2002 laying down detailed rules for the implementation of the Financial Regulation

- The **exclusion** and **selection** criteria applying to the contract, save in case of a restricted or negotiated procedure following publication of a notice. In such cases those criteria shall appear solely in the contract notice or the call for expressions of interest;
- The award criteria and their relative weighting.
- **Exclusion and selection** criteria: They have to be non-discriminatory. The following selection criteria shall apply:
 - a) Exclusion criteria (see art. 93 and 94 of the Financial Regulation, and art. 134 of the Implementing Rules of the Financial Regulation) concern the eligibility of the service provider to take part in the procedure by verifying the possible grounds for exclusion referred to in the Financial Regulation (e.g., if the service provider is bankrupt);
 - b) Selection criteria (see art. 135, 136 and 137 of the Implementing Rules of the Financial Regulation): criteria for assessing his financial, economic, technical and professional capacity⁴⁶.
- **Award** criteria (see art. 138 of the Implementing Rules of the Financial Regulation): there are two possible award procedures:
 - a) Automatic award procedure, in which case the contract is awarded to the tender which, satisfying all the conditions, quotes the lowest price;
 - b) The best value-for-money procedure. If this procedure is chosen, the Implementing rules specify that the tender offering the best value-for-money shall be the one with the **best price-quality ratio**, taking into account criteria justified by the subject of the contract such as the technical merit and the price quoted. The weighting applied to each criterion for determining the best value-for-money shall be indicated. The weighting applied to price in relation to the other criteria must not result in the neutralisation of price in the choice of the contractor.

The exclusion, selection and the award criteria shall be clearly defined and separate. The issue of independence (conflict of interest) can be dealt with as a condition for eligibility (exclusion criteria) in accordance with article 94 of the Financial Regulation.

As already stated above, the methodological quality of the proposal is often linked to the overall budget for the evaluation project. Furthermore, the quality of the method depends on its capacity to provide answers to main evaluation questions. The following list gives some indication on how to tackle the assessment of methodological quality:

The proposal of the evaluation team:

- Does it include the collection of relevant information?
- Is it based on rigorous analytical techniques?
- Is it able to clarify the evaluation criteria in an impartial manner?
- Is it likely to produce credible findings?

The following example from DG INFSO (intermediate evaluation of Internet Action Plan) illustrates possible elements of award criteria.

⁴⁶ The profile of individual members of the evaluation team falls under the category of technical capacity

<u>Example – Award criteria.</u>	
EVALUATION CRITERIA	AWARD CRITERIA
Understanding (30%) ⁴⁷	<ul style="list-style-type: none"> • Understanding of objectives and tasks
Methodology (40%)	<ul style="list-style-type: none"> • Credibility and rigour of evaluation design, • Credibility of data collection techniques, • Credibility of data analysis techniques.
Project management (30%)	<ul style="list-style-type: none"> • Clarity of objectives and milestones, • Soundness of work method, • Soundness of resource allocation.
Price	<ul style="list-style-type: none"> • Professional fees, • Travel and subsistence costs, • Translation expenses, • Other expenses.

Concerning procurement of evaluation services, DG BUDG has established a Framework Contract⁴⁸ and an open Call for manifestation of interest⁴⁹ ('AMI-list') in the field of evaluation studies and other evaluation services. These instruments are at the disposal of all Commission services and they aim to facilitate the procurement of evaluation services, in particular by reducing the time and administrative steps required for selecting an external consultant and establishing a contract.

3.3.8 The structure of the final report

The Terms of Reference should also specify the number of reports and their contents and format. For more on the structure of the final report, see Section 5.1, where the reporting is presented. As far as a good example of the structure of an evaluation report is concerned, see the Annex 3 of the Evaluation Standards.

⁴⁷ Please note that the relative weights (percentages) given are specific to the example. These weights may vary between evaluation projects depending on number of criteria, such as the complexity of the intervention to be evaluated, its geographical and financial scope etc

⁴⁸ <http://www.cc.cec/budg/networks/i/evaluation/en/framework/index.htm>

⁴⁹ <http://www.cc.cec/budg/networks/i/evaluation/en/list-experts/expertslist.htm>

4 CONDUCTING EVALUATIONS

This chapter addresses the different arrangements related to conducting evaluations. It will first discuss on how to set up a steering committee for an evaluation project. After that the text moves to actual carrying out of evaluation, including the administrative set-up and methods to be employed. The chapter will close with the validation of the evaluation report.

Standards

C.5: The evaluation function shall decide on the advisability of setting up a steering group for each evaluation.

C.6: The steering group shall be set up and perform its tasks in compliance with the rules on conflicts of interest and the directives on public contracts.

C.7: The tasks of the steering group shall be, as a minimum, to:

- a. facilitate the evaluator's access to the information needed to perform his/her work,**
- b. support the evaluation work, particularly from the methodological standpoint,**
- c. take part in assessing the quality of the evaluation at the appropriate juncture.**

Good practice C.2: For external evaluations, it is advisable to set up a steering group including the departments most concerned, notably a representative of the operational unit responsible for managing the area evaluated and a representative of the evaluation function.

4.1 Steering Groups

This section deals with the question of how to set up a steering group. The first sub-section addresses the question of how to compose a steering group while the second one examines the tasks and responsibilities of such a group.

4.1.1 Composition of the Steering Groups

The Steering Group is usually composed of a group of officials, from the Commission, who are directly concerned and/or involved with the intervention subject to the evaluation and furthermore have some knowledge of evaluation.⁵⁰

Depending on the nature of the intervention and the expectations and needs stemming from the decision-making process the steering group can, and quite often is, broadened to include also members outside the Commission. This was the case for 20% of the Steering Groups in the Commission in the period 1998-2002. In cases where the Steering Group is broadened to include also members outside the Commission, a Commission official should normally chair the work.

⁵⁰ It is worth noting, however, that in the Commission context a Steering Group differs from a committee for opening the tenders and a committee for evaluation of tenders. The latter are always nominated by the authorising officer as stipulated in articles 145 and 146 of the Implementing rules of the Financial Regulation. However, in normal practice, a number of the steering group members will also participate in the tender opening and evaluation of tenders committees.

The advantages of a broadly based Steering Group can be presented as follows:

- Easier access to information and a better understanding of the facts and events which are taking/took place while the implementation of the intervention,
- Interpretation and recommendations which take into account all the important points of view,
- Better and quicker dissemination of conclusions and recommendations,
- Better acceptance of the evaluation and its findings by the stakeholders.

Different types of stakeholders may be considered for the Steering Group:

- Actors responsible for the implementation including managers, operators and agencies, and
- The groups affected by the intervention, including direct and also indirect addressees.

The involvement of political authorities may also be considered, while, however, taking the different roles of the Institutions into account.

The different members of Steering Groups may, at least partly have different roles to play. Some members may, according to their expertise or organisational background focus on particular aspects such as provision of data or commenting and validating the results drawn up from the evaluation process. In certain cases, in particular where actions with devolved implementation are evaluated centrally, there may be a case for consulting established committees as a complement to the work of the steering group (e.g., concerning dissemination).

4.1.2. Tasks of the Steering Group

The Steering Group normally has a role to play before, during and after the execution of the evaluation project. Firstly, it transforms the political issue/questions into operational evaluation questions. Secondly it supervises and guides the evaluation project in order to ensure that the selected evaluation team carries out and produces the actual evaluation report in line with the strategy taken in the Terms of Reference. It will normally also have a role to play in the follow-up to the evaluation, for example concerning dissemination or in order to prioritise or operationalize the recommendations and devise an action plan.

The role of the Steering Group can also be expressed in terms of the meetings normally taking place during an evaluation process. The Steering Group would then be involved in at least meetings on the (i) terms of reference, (ii) inception report, (iii) progress report(s), and (iv) final report(s).

The Steering Group should hence normally contribute to the following steps:

- Prepare the Terms of Reference,
- Provide access to information,
- Support and monitor the work of the evaluator,
- Assess the quality of the final report,
- Dissemination of the results and other follow-up.

4.2 Carrying out evaluations

4.2.1 Administrative set-up

The choice between in-house work and entrusting the whole or parts of the evaluation to outside sources depends on a number of factors, but often the key determinants are the overall purpose of the evaluation exercise and time and resources available for it.

The use of external contractors is particularly suitable, when the evaluation has a strong summative dimension (i.e. evaluation provides information with the objective of transparency and accountability purposes). On the other hand the in-house choice is useful when the summative dimension is weak, but the formative dimension (i.e. objective of the evaluation is to increase organisational learning and improvement in the quality of the intervention) is strong. In practice this makes ex ante evaluation more likely to be carried out internally than other types of evaluations. However, in reality an evaluation project will often combine both summative and formative aspects.

The other important variable influencing the administrative set-up is the level of resources available for the evaluation project. If the time available for conducting the evaluation is very short, in-house work is often the only possible solution available (e.g. because of the time necessary to respect the procurement rules). Again, this is particularly the case with the ex ante evaluation, where the assessment is carried out in close interaction with the design of the intervention, and time available for conducting the evaluation may be limited (furthermore, in the case of ex ante evaluation, the direct feedback of findings into the design of the intervention is important, a process facilitated by undertaking the work in-house). Nevertheless, even in cases like this, parts of the assessment, such as carrying out surveys or organising workshops, can be entrusted to the outside contractor.

The relative advantages of external and internal evaluations can be presented as follows:

External evaluation / Advantages:

- Strengthens the credibility of the evaluation project from an accountability perspective by underpinning the independence, objectivity and transparency of the process,
- Optimises use of scarce human resources and brings in expertise, which might not be available in the Commission.

Internal evaluation / Advantages:

- Supports directly organisational learning,
- Mobilises available in-house knowledge,
- Facilitates direct feedback into the design of an intervention,
- Can be carried out whenever financial resources are limited

Sometimes, however, it may be advantageous to combine elements of internal and external evaluation within a single evaluation project. This is especially the case in certain types of ex ante evaluations (c.f. 1.1.3).

In the Commission, the use of external experts has been dominant. Almost 90% of the evaluations since 1996 were conducted externally. However, a slight rise in the share of evaluations conducted internally can be observed in the more recent period.

4.2.2 Implementing the evaluation

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 for the time, budgetary and data constraints of the project. In addition, the methodology of evaluation will include an element of “desk research”, in other words, a process by which the evaluator collates and analyses data from secondary sources (programming and policy documents, monitoring reports and data, statistical data, relevant studies and evaluations, etc.), particularly at the outset of the evaluation and, later on in parallel to the collection of data in the field.

Evaluation tools and techniques are implemented to fulfil one or more of the following functions:

- to help structure an evaluation.
- to collect of data in/outside the Commission.
- to analyse data.
- to aid the formulation of judgements.

It will be useful, already in the Terms of Reference, to distinguish the following four phases of evaluation. The **structuring phase** (cf. section 3.2) of an evaluation covers the period from the drawing up of the terms of reference by officials to the delivery of the inception report by the evaluator. In this phase the effects to be evaluated are chosen and classified and the observation tools are clearly defined. The structuring phase embraces preparatory work, drafting the terms of reference and it normally culminates with the inception report. Obviously, the inception phase must not change the focus of the work as set out in the terms of reference, but it can make it more specific and operational. The purpose of the structuring phase is to ensure that evaluation is focused so as to provide the information sought by its users (e.g. through the identification of evaluation questions, success criteria, indicators, target levels and data sources) and of ensuring that the evaluation method to be implemented is adapted to the data collection and analysis tasks ahead and will provide a clear basis for judging the performance of the intervention. Structuring tools may therefore be used by officials to support the elaboration of the TOR as well as by the evaluator to facilitate the understanding of an intervention, establish an analytical framework, etc. The structuring phase normally culminates with the inception report which proposes to the steering group how to make the evaluation strategy, as set out in the Terms of Reference, operational.

The **data collection phase** in the field commences after the inception report has been delivered, on the basis of a finalised evaluation method presented by the evaluator and approved by the steering group. Data collected by the evaluator will usually be of both a qualitative and quantitative nature, and will provide factual information along with perceptions and opinions.

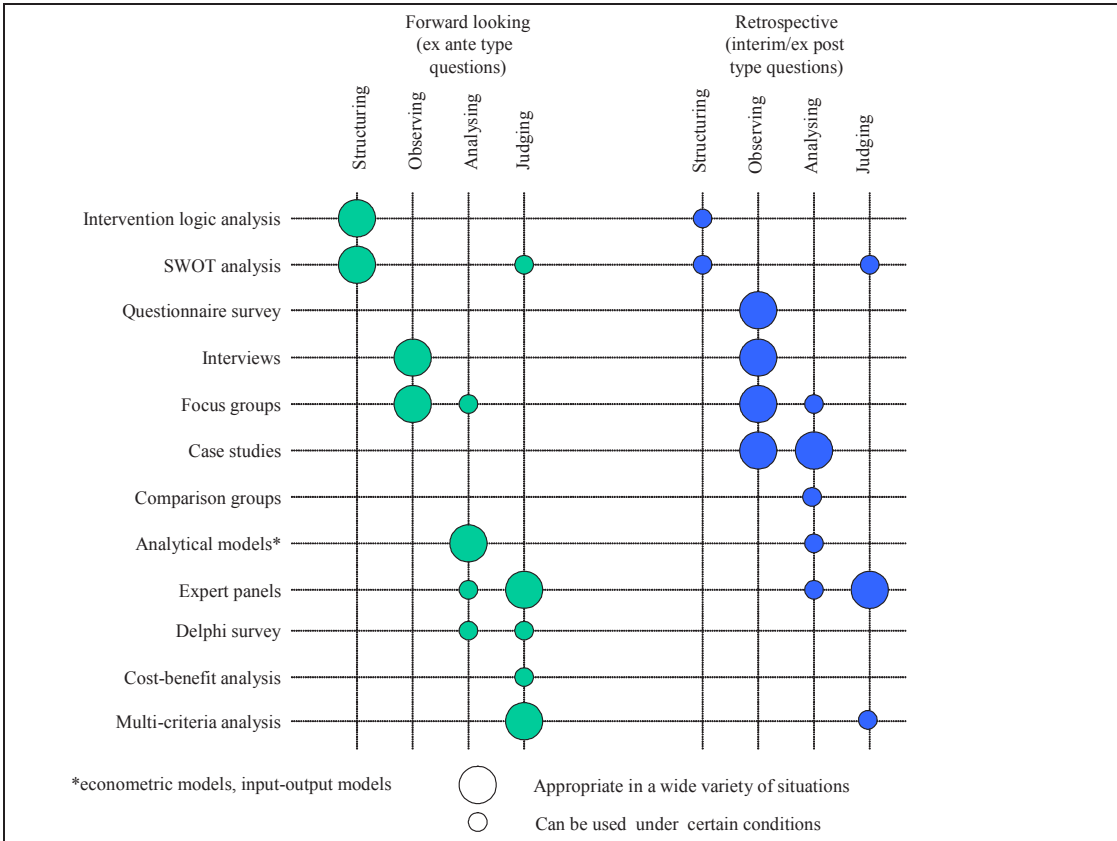
The **analysis of data** firstly takes place at the level of individual data collection tools and techniques (e.g. the analysis of a survey, interviews, etc.). Secondly, an analysis will take place across data collected, using a number of tools and techniques, and from different informants and documentary, statistical sources and monitoring systems. This *cross-analysis* will occur when data collection and initial analyses are completed. An essential part of work of the evaluator at this point is an analysis of causality and the attribution of effects to the intervention.

The collection of data in the field can be particularly time consuming (in terms of human resource requirements) and requires skills, as is the case with the analysis of data. In many cases these reasons contribute strongly to the choice of conducting external evaluations.

The final stage of the evaluation consists of the **formulation of judgements**. Once all the data has been analysed and the evaluation results compiled the success of an intervention can be judged, a process which will focus in particular assessing an intervention in respect to specific evaluation questions that reflect generic evaluation issues such as effectiveness, efficiency, utility, sustainability and relevance (cf. section 3.2).

The tools and techniques that will be used during the evaluation also depend on whether the evaluation questions to be addressed are forward looking or retrospective in nature since some tools and techniques may be better suited to one or the other of these situations.

An overview of some of the tools and techniques that are likely to be used by evaluators in different evaluation situations in a Commission context are presented in the box below and a more detailed description of their key characteristics figures in Annexe E.



Source: Adapted from MEANS.

4.2.3 Validation of evaluation reports

An evaluator will deliver a number of different reports at various key stages of the evaluation process: inception report, intermediate report(s), draft final report and final report. Each report should be critically assessed as it provides a basis for tracking quality of the work done by the evaluator. A structured quality assessment framework exists for the draft final and final reports.

Inception report. This report draws to a close the structuring phase of an evaluation. It describes in detail how the method proposed by the evaluator is going to be implemented in the light of an examination of the quality and appropriateness of existing data, and in particular how the method will answer each evaluation question and provide a judgement. It may also contain the results of some initial analyses, for example in the form of typologies of actions, addressees, countries, etc. This document provides an opportunity to make a final check of the feasibility of the method proposed and the extent to which it corresponds with the information needs outlined in the terms of reference, before the evaluation gets fully underway. The meeting on the inception report is crucial, because it indicates the extent to which the contractor has understood his/her tasks, and moreover, it still offers the Commission the possibility to give final instructions based on the preparatory work during the first weeks or months, and correct any possible misunderstandings.

Interim report. This report will provide information about initial analyses of data collected in the field (primary data) and secondary data. Much of the required data may have been collected at the stage of the evaluation when this report is delivered and the evaluator may already be in a position to provide preliminary answers to some of the evaluation questions. This report provides the evaluation manager and the steering group with the opportunity to check whether the evaluation is on schedule and whether the evaluation has actually focused on the specified information needs. It also provides the basis for a dialogue between the evaluator and the steering group about the adequacy of analyses, the factual accuracy of observations and the realism of assertions and interpretations.

Draft final report. This document will provide the conclusions of the evaluator in respect to the evaluation questions in the terms of reference. These conclusions will be clearly based on evidence generated through the evaluation. Judgements provided should be clear and explicit. The draft final report may also contain some exploratory recommendations developed on the basis of the conclusions reached by the evaluator. The draft final report should be subject to a preliminary quality assessment that will verify the extent to which it respects the criteria presented in the Evaluation Standards, as shown in the box below. Concerning the quality assessment, there are already quality assessment frameworks in the Commission, such as the grid presented in Annex G.

Standards

D.1: The substance of the evaluation reports shall be relevant, based on rigorous analysis, meet the quality criteria laid down in the specifications (cf. example in annex 2 to the standards) and comply with the deadlines

D.2: The evaluation reports shall describe the purpose of the evaluation and its context and also the objectives, questions, procedures, results and reasoned conclusions of the evaluation, so as to make available the essential information in an easily understandable form.

D.3: The report shall describe the information sources in such detail that the correctness of the information can be assessed. The data collected or selected shall be adapted to the methodologies used and be sufficiently reliable for the expected use.

D.4: The prospects and reasoning on which interpretation of the results is based shall be described and explained. The results should follow on logically and be substantiated by data analysis and interpretations based on carefully-presented explanatory hypotheses.

D.5: The final evaluation report shall present the results and conclusions of the evaluator and the tenor thereof shall not be amended without his/her agreement.

D.6: The conclusions and any recommendations shall be rigorous and not distorted by personal or partisan considerations. The recommendations shall be comprehensible, useful, applicable and detailed enough to be brought into effect.

Final report. It will take into account the results of quality assessment and discussions with the steering group about the draft final report insofar as they do not interfere with the autonomy of the evaluator in respect to his or her conclusions.

Validation of the evaluation report is the final part of the chain of quality assurance, a process that already starts with the preparatory work for drawing up the Terms of Reference. The quality assessment is an important step in the validation, and should not await the final report. If a preliminary quality assessment is carried out on the draft final report, it will help the evaluation team to focus the subsequent work leading up to the production of the definitive final report.

As an example of possible phases in the validation process, consider the following example from DG AGRI.

Example – Validation of an evaluation report

I. DRAFT FINAL REPORT PHASE

1. Steering Group internal meeting

- the idea is to reach a common position of the steering group before the meeting with the contractor takes place. The group goes through question by question the draft report.
- after the meeting, the evaluation chairperson/manager finalises the common global assessment of the submitted draft report.

2. Steering Group meeting with the contractor

- after a brief discussion on general remarks, the meeting concentrates on specific topics requiring in-depth discussion, such as data collection and analysis and conclusions.
- the requested improvements should be of methodological nature, and should not aim at distorting the contractor's judgement. Discussions about conclusions are possible, if based on data, analyses, findings and judgement criteria.

II. FINAL REPORT PHASE

3. First version of the final report and quality assessment meetings

- critical assessment concentrates on global quality, and when necessary, minor factual details. As a result of the global assessment, the report is either i.) good enough to be edited as such, or ii.) it still contains errors, which the contractor has to take into account before final editing.

4. Final Steering Group internal meeting on the definitive final report

- steering group examines the consultant's conclusions and recommendations in terms of methodological validity only.
- the quality assessment is discussed, substantiated and revised as necessary.

5 FINAL REPORTING, DISSEMINATION AND USE OF EVALUATION RESULTS

This chapter addresses the feedback mechanisms related to evaluation findings at the end of the evaluation project (the reports to be submitted during the evaluation process are dealt with in section 4.2.3). These mechanisms relate to both reporting and dissemination and use of evaluation results. The final reporting takes place when the contractor transmits the evaluation report to the Commissioning body/Steering group. Dissemination, on the other hand, refers to the set of activities by which an evaluation's findings are made available to the target group as well as to the general public.

Although reporting and dissemination are the final steps of the evaluation project, it is important that the evaluation function and/or the Steering group thinks about strategies for communicating the evaluation results already at the preliminary stages of the evaluation project.

Standards.

B.7: Appropriate feedback mechanisms shall be provided so that all types of evaluation results are transmitted effectively to all persons responsible for decision-making (senior management and stakeholders, including those responsible for policies).

B.8: The Directorate-General and Services shall report on the results of the evaluation of their activities in their Annual Activity Reports.

B.9: Feedback mechanisms shall contribute, on the basis of evaluation results, to the formulation, planning and revision of policies, programmes and activities.

B.10: The operational departments shall examine the evaluation conclusions and possible recommendations and describe the action they propose to take.

B.11: The DGs and services shall transmit to the Budget DG and to the SG copies of all the final evaluation reports.

B.12: Save in duly substantiated cases of confidentiality, evaluation results shall be made publicly available. The means of communication shall be readily identifiable and accessible.

B.13: The evaluation results shall be transmitted in an appropriate form to meet the needs of decision-makers, beneficiaries and citizens and to fulfil the obligations for reporting in accordance with the legal bases.

5.1 Reporting

Standard D.2

The evaluation reports shall describe the purpose of the evaluation and its context and also the objectives, questions, procedures, results and reasoned conclusions of the evaluation, so as to make available the essential information in an easily understandable form.

The following text on reporting relates mainly to interim and ex post evaluations. As far as **ex ante evaluation reports** are concerned, it is recommended to provide a single unified final report, or at minimum, a summary of the key findings of the report to respect the requirements

of the article 28 of the Financial Regulation⁵¹. However, the results of an ex ante evaluation should always be presented in an abridged form in the context of the legislative financial statement (see section 5.2 below).

Since the evaluation report is normally the concrete and visible product of the whole evaluation project, it is important that it is well written and presented. The style and presentation contribute to the structure and clarity of the report, and both of these are thus addressed below.

The text below focuses on interim and ex post evaluation reports. As far as ex ante evaluation is concerned, its structure should always be adapted to its specific purposes. As an example of a possible structure for ex ante evaluation report, see Annex H.

The structure of the evaluation report should reflect the different uses and expectations of the final report. A broad classification into following three parts is usually recommended⁵²:

- **Executive summary.** This part is destined primarily for the higher levels of the hierarchy and Cabinet and provides a short synthesis (usually, maximum of five pages) of the main conclusions of the evaluation, the key points of evidence underpinning them and the resulting recommendations.
- **Main report.** It is aimed at those directly or indirectly involved in the evaluation and the management of the activity or an intervention. The main report presents in full the results of analyses, conclusions and recommendations of the evaluation. It will also contain a description of the activity evaluated and the context of the evaluation as well as the methodology used and its main limitations.
- **Technical annexes.** One or more documents will collate the technical details of the evaluation. Intermediate deliverables, for example case study monographs, will also be presented where appropriate in the technical annexes. Typically, the main users of the annexes are operational services managing the intervention, other evaluators or academics.

The main report will contain three different types of chapters: introductory, descriptive and substantive. While each chapter will present essential information to the reader, the main focus of the document is on evaluation's results, their interpretation and subsequent conclusions and recommendations (i.e. the substantive chapters). These chapters will consequently represent the biggest share of the report's contents.

Introductory chapters on evaluation objectives and scope. The introductory chapters will enable the reader to understand the context in which the evaluation took place, in particular, why the evaluation was conducted, why at this particular moment in time, and how the information it has produced will be used and by whom.

The focus of the evaluation will be clearly presented to the reader in terms of the substantive questions it has answered and the time period and geographical scope for which the answers are relevant. The breadth of the work in terms of the "taking into account" of related public interventions will also be clearly announced.

Descriptive chapters presenting the activity to be evaluated and the method used. Chapters of a descriptive nature will present to the reader:

- The design and implementation of the activity, and the context in which it is being or has been implemented

⁵¹ 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).

⁵² Adapted from the Guide for the evaluation of the activities of DG ADMIN, March 2002

- The evaluation method and its limits, in particular:

How the method answered the evaluation questions;

What primary data (from interviews, questionnaire survey, etc.) and secondary data (from existing studies and evaluations, monitoring systems, statistical sources, etc.) have been collected and used in analyses;

Analyses conducted (in particular the extent to which questions of causality and attribution have been addressed).

Given that the main report should not be a voluminous document (no more than 100 pages in length is an often used “rule-of-thumb”), it will contain fairly succinct descriptions of the activity and the evaluation method (i.e. enabling the average reader to understand what is being evaluated and how). A more detailed description will be presented in the appendix to the main report, with appropriate cross referencing of documents.

Substantive chapters presenting the results of analyses, their interpretation and subsequent conclusions. At the heart of all evaluation reports are the substantive chapters presenting the evaluation results, conclusions, recommendations and/or lessons learnt.

Results will be presented to the reader in a clear and understandable way (for example using tables, graphs and bullet points for clarity when presenting key points). Results will be fully explained while at the same time being kept clearly separate from the evaluator’s interpretations/judgements. The way in which results are presented should allow the reader to be able to trace back to the analyses conducted during the evaluation and the data used.

When reading these chapters the user will be able to follow the lines of reasoning that take the evaluator from the results of analyses through to his or her conclusions, so as to be able to accept them, challenge them, or point out where they require strengthening. Conclusions will directly give rise to specific recommendations or to broader lessons learnt, as appropriate.

Apart from a clear structure of the final report, it is also important that the text of the final evaluation report is understandable and clear, and that the findings, conclusions and recommendations of the evaluation follow a logical flow.

The following list is a reminder of the topics that any potential reader must be able to understand when reading the final report.

- The purpose of the evaluation,
- What was evaluated,
- How the evaluation was designed and conducted,
- What evidence was found,
- What conclusions were drawn,
- What recommendations were made and/or lessons learnt.

5.2 Dissemination and disclosure of evaluation results

Financial Regulation, art. 28 (1):

Any proposal submitted to the legislative authority which may have an impact on the budget, including changes in the number of posts, must be accompanied by a financial statement.

The disclosure standard, B.12, is underpinned by a recent Regulation⁵² implementing Article 255 of the Amsterdam Treaty. This Article grants a right of access to European Parliament, Council and Commission documents to any Union citizen and to any natural or legal person residing in a Member State. The definition of a "document" in the Regulation is a broad one: no category of document is excluded *a priori* from the right of access, including classified documents, internal documents, and documents originating with third parties. Refusal to grant access must be based on one of the exceptions provided for in article 4 of the Regulation. For the Commission, the Regulation is interpreted by a set of internal rules⁵³, a staff guide⁵⁴ produced by the Secretariat-General, and a "Citizen's Guide"⁵⁵. In an evaluation context, the essential point to bear in mind is that the onus is on the services to justify why access to a particular evaluation report should be denied.

Turning to dissemination, it is important to note that the knowledge generated from an evaluation process does not automatically reach the intended and potential users. Hence, often it is useful to design a specific dissemination strategy whereby all the key elements, which influence the evaluation being used to support the management and decision-making processes are identified and arrangements for their implementation envisaged.

The first step in a dissemination strategy is to identify the key potential users at the outset of the evaluation and present them already in the Terms of Reference. This enables the evaluators to know right from the beginning who these potential users are.

Below is a short list of potential evaluation users:

- **Key policy-makers and interested institutional parties.** In the case of EU-activities, this group includes the Cabinet or possibly College, the EU Parliament and its various Committees, the Council, the Court of Auditors as well as the Member States. As far as individual DGs are concerned, a separate copy of the final evaluation report shall always be sent to DG Budget and SG. DG Budget needs these reports on a regular basis in order to comply with the obligations vis-à-vis the Budgetary Authority⁵⁶.
- **Managers and operators of the intervention being evaluated.** This group may consist of individual DGs, agencies and other public bodies involved in the implementation of the

⁵² Regulation 1049/2001 of the European Parliament and the Council, 30 May 2001
http://europa.europa.eu.int/comm/secretariat_general/sgc/acc_doc/docs/1049EN.pdf

⁵³ Commission Decision 2001/ 937, 5 December 2001
http://europa.eu.int/eur-lex/pri/en/oj/dat/2001/l_345/l_34520011229en00940098.pdf

⁵⁴ Staff Guide to Public Access to Commission Documents
http://www.cc.cec/home/dgserv/sg/docinter/guide_en.htm

⁵⁵ Access to European Commission Documents: A Citizen's Guide
http://www.europa.eu.int/comm/secretariat_general/sgc/acc_doc/docs/en.pdf

⁵⁶ Since the 2001 discharge procedure, DG Budget has been committed to make available frequently synthesis reports on finalised evaluations. Moreover, the introduction to part B of each budget states that: "there must be regular assessment of implementation of all programmes and projects and grants to organisations financed or co-financed by the operating part of the European Union budget. The budgetary authority must be informed about the results of those assessments, even if this is not expressly provided for in the legal basis".

intervention. Especially for devolved interventions there will also be operators in the Member States or, delegations in third countries etc.

- **Addressees of the intervention.** Bodies which are affected either positively or negatively, directly or indirectly by the intervention. Examples of addressees are NGOs, private enterprises, individual persons etc..
- **Other Services.** Commission services can assist one another by making efforts to share evaluation findings. Often, an evaluation in one sector will produce findings which are relevant and interesting in another.
- **Other interest groups.** These are, for example, organisations, groups or individuals having a general interest in the intervention being evaluated as well as academics with a specific scientific interest.

Together with identifying the potential users of the evaluation, it is often useful to analyse their characteristics in order to develop appropriate communication vehicles. The following list offers some basic questions, which can be used when analysing the target audience of evaluation⁵⁷.

- How is the target audience composed?
- What do they need?
- What is their knowledge of the evaluation?
- Were they involved in the evaluation design? If so, to what extent? If not, why?
- How might they be encouraged to attend any presentation?
- What advantages and disadvantages might result to them from the evaluation?
- Which evaluation questions are of interest to them?
- What other issues are important to them?
- Are they likely to object to particular findings, conclusions or recommendations and how these objections might be overcome?
- How interested will they be in the fine details compared to the overall picture?

5.3 Channels for diffusing evaluation findings

The final step in the dissemination strategy is to explore different sorts of communication vehicles that may be used, when reporting on the results of the evaluation. Apart from the more traditional paper documents, there exists nowadays a wide range of other, supplementary communication means. It is worth noting that the choice of the communication means is often linked to the targeted audience, as shown below.

⁵⁷ Adapted from the Guide on ex post and intermediate evaluation, January 1997.

Different ways of diffusing evaluation findings to different audiences⁵⁸:

	Decision-makers	Managers	Addressees	Other interest groups
Meetings		X		
Draft report		X		
Confidential note	X ⁵⁹	X		
Public report	X	X	X	X
Synthesis note⁶⁰	X	X		
Presentation	X	X		X
Brochure/ article			X	X
Press conference / release			X	X
Communication	X	X		
APS-decision	X	X		
Preliminary Draft Budget / Activity Statements	X	X		
Annual Activity Report	X	X		
Fiche contradictoire / Action Plan		X		
Explanatory memorandum + legislative financial statement	X	X		X
Commission's Annual Evaluation Review	X	X		X

⁵⁸ Adapted from the Guide for the evaluation of the activities of DG ADMIN, March 2002.

⁵⁹ E.g. in the case of Cabinets

⁶⁰ For example, an Executive summary or an internal note translating the conclusions and recommendations of evaluation report into the actual setting of the operational services.

As a minimum, the evaluation results should be reported in the Commission's Annual Evaluation Review. Moreover, as stated in the article 28 of the Financial Regulation, the results of ex ante and ex post evaluations should also be reported in the legislative Financial Statement accompanying a proposal submitted to the legislative authority. These results should also be indicated in the explanatory memorandum related to the proposal.

Example – Use of evaluation information in the sections of the Legislative Financial Statement

The results of the ex ante evaluation should be integrated into the various sections of the Legislative Financial Statement. The latter should also explain how and when the ex ante evaluation was conducted, in which form it is available and explain how the results of the evaluation have been integrated into the proposal in question.

An ex ante evaluation cannot simply be replaced by a previous interim or ex post evaluation. However, such evaluations can, if the conditions are right, be used as a basis or as the main input to an ex ante evaluation. In such cases it will be necessary to provide additional information on:

- the need to continue with such intervention;
- the validity of the objectives and the setting of results/effects to be achieved in the new period;
- a verification of the arrangement for implementing the intervention and the expected level of resources;
- if and how lessons learned from the previous interim or ex post evaluation have been taken into account in the new proposal.

Finally, the article 33 of the Financial Regulation also calls for evaluation results when the budget is established by stating that evaluation results shall be consulted and referred to as evidence of the likely merits of a proposed budget amendment.

Moreover, key elements should be taken up in the relevant ABM documents, such as APS decision, Annual Management Plan and Annual Activity Report. The latter is important because the Directors-General and Heads of Service sign a report on the achievements of their service. Evaluation results help to inform a service on the actual impact of Commission activity, and enable adjustments to be made. If appropriate, an adjustment can be made to objectives fixed for an Activity in the following Annual Management Plan.

In order to ensure an efficient dissemination of evaluation results, it is strongly recommended that whenever the technical facilities allow it, the reports should always be published in the intranet/internet of the DG or unit.

5.4 Use of evaluation results

It is important to support the process of transposing evaluation results into a form that is most likely to lead directly to use. There may be an agreed policy to follow up on the evaluations to maximise the probability that they feed back into the decision making process at whatever level is appropriate or the transposing may take a form of a specific action plan. The box below contains some examples of good practice in promoting instrumental evaluation use:

Example – promotion of evaluation use

- The use by DG AIDCO evaluation unit of what is called a “fiche contradictoire”, a document which presents the recommendations of an evaluation to the relevant operational unit and invites the latter to accept or to refuse explicitly the recommendations of the evaluator.
- The assessment by the DG AGRI evaluation unit of the use of an evaluation six months after it was reported on as part of standardised follow-up procedures.

Reports or Communications of the Commission can also be used as a vehicle for capitalising on evaluation results. DG INFSO offers a recent example covering the final evaluations of the INFO 2000 and MLIS programmes.

Example – Communication as a tool for capitalising evaluation results

After the final evaluation reports had been published, DG INFSO prepared a Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions. This Communication⁶¹ lists the main recommendations of the evaluation reports together with the comments and actions to be taken up from the Commission in formulating a follow-on programme proposal. Moreover, if the Commission disagrees with any recommendation of the evaluation report, a reason for that is given.

Apart from the above follow up processes, there are also other factors, which influence the way evaluation results are used. One such factor is the degree to which potential users are involved in designing and focusing the evaluation. The aim of involving multiple potential users is not only to ensure that various interests are being taken into account, but also to make these users aware of future use and dissemination of the evaluation.

Another factor influencing the use of evaluation results is the needs and expectations which potential users have towards evaluation. The information needs typically vary, depending on the overall purposes of evaluation (as specified in the section 1.1.1).

If the objective of the evaluation is only to increase the quality of intervention, then the range of potential user types is often quite narrow (such as managers and direct beneficiaries). On the other hand, if the evaluation aims to provide input for setting political priorities and related allocation of resources, the range of potential users becomes more diverse, including also key decision-makers. In order to ensure a proper use of evaluation, the key aspect is to target (the often multiple nature of) information stemming from evaluation products to the right audience, as described above.

⁶¹ COM(2001) 276:
http://europa.eu.int/information_society/programmes/evaluation/pdf/cominfo2000mlis_en.pdf

ANNEX A: EVALUATION STANDARDS

(in the parentheses, the Chapter/Section where the standard is addressed in this document):

A. Profile, role, tasks and resources of the evaluation function

1. The evaluation function shall be clearly visible in the DG's structure, the organisation chart or *Guide des Services* giving details of the unit, sector, coordinator, network members and person(s) responsible. (See section 2.1)
2. Evaluation shall be clearly identified in the list of each DG's operations so that the human and financial resources set aside for evaluation can be easily identified. (See section 2.1)
3. Each DG or Service shall define the tasks, responsibilities, organisation and procedures for running, consulting and informing the evaluation function. (See section 2.1)
4. The role of the evaluation function shall be clearly defined and distinguished from the roles of other functions such as monitoring, control and audit. (See section 1.1.2)
5. Through the evaluations, the evaluation function provides support in the formulation of policies and programmes and also in management processes. (See section 2.1)
6. To fulfil this role, the evaluation function shall, as a minimum. (See section 2.2):
 - a. co-ordinate the evaluation activities of the DG or Service,
 - b. help to anticipate decision-making requirements in the fields covered by the DG or Service,
 - c. establish a multiannual evaluation programme and an annual evaluation plan for the DG or Service consistent with decision-making requirements and accordingly 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,
 - d. co-ordinate and monitor implementation of this programme and this plan at central level; where appropriate, it shall assist in decentralised implementation of this programme and this plan,
 - e. define quality standards for evaluation activities on the basis of these standards and, where appropriate, its own quality requirements, more specific and detailed or better adapted to its policy area,
 - f. promote the methodological quality and consistency of the evaluation activities of the DG or Service,
 - g. check on how the evaluation recommendations have been taken into account,
 - h. ensure that the evaluation results can be useful in defining departmental priorities and in improving the quality of their policies, programmes and activities in such a way that these results can also be used in defining the Commission's political priorities and budgetary decisions.
7. The human and financial resources available for evaluation work shall be clearly identified and planned. These resources shall be sufficient for implementing the

multiannual evaluation programme and annual evaluation plan of the DG or Service.
(See section 2.3)

8. In relation to the evaluation plans approved, the resources earmarked for the evaluation function shall be adequate and appropriate in terms of funds, staff and skills so as to ensure that they can fulfil their responsibilities effectively.
(See section 2.3)

B. Management of evaluation activities

1. The DGs shall plan their evaluation activities in accordance with (See section 2.2.2):
 - a. the Financial Regulation and its implementing provisions;
 - b. the Communication on evaluation of 26 July 2000 (SEC(2001) 1051) extending the principle of evaluation to all activities in the context of activity-based management and budgeting (ABM/ABB);
 - c. the rules on evaluation in the Communication on implementation of activity-based management (ABM) of 25 July 2001 (SEC(2001) 1197/6&7)
 - d. the Commission's rules on impact assessment
2. This planning shall take the form of a multiannual evaluation programme and an annual evaluation plan. (See section 2.2.2)
3. The multiannual evaluation programme shall be drawn up on the basis of the life cycle of the policies and programmes and in accordance with the nature of the activities of the DG or Service. The programme shall be approved by the Director-General –or, if the specific decision is taken, the Commission- and reflects future political priorities and operational needs. (See section 2.2.2)
4. The multiannual evaluation programme shall spell out the division of tasks between the evaluation function and the operational departments in the centralised or decentralised implementation of the evaluation activities. (See section 2.2.2)
5. The annual evaluation plan shall be drawn up by the evaluation function and used to implement the multiannual evaluation programme. The plan shall comprise an estimate of the human and financial resources needed to bring it into effect. The DG or Service shall adopt and supplement the annual evaluation plan in its annual management plan (AMP). (See section 2.2.2)
6. The audit and evaluation plans shall be clearly separated. (See section 2.2.2)
7. Appropriate feedback mechanisms shall be provided so that all types of evaluation results are transmitted effectively to all persons responsible for decision-making (senior management and stakeholders, including those responsible for policies). (See chapter 5)
8. The Directorates-General and Services shall report on the results of the evaluation of their activities in their Annual Activity Reports. (See chapter 5)
9. Feedback mechanisms shall contribute, on the basis of evaluation results, to the formulation, planning and revision of policies, programmes and activities. (See chapter 5)

10. The operational departments shall examine the evaluation conclusions and possible recommendations and describe the action they propose to take. *(See chapter 5)*
11. The DGs and Services shall transmit to the Budget DG and to the SG copies of all the final evaluation reports. *(See chapter 5)*
12. Save in duly substantiated cases of confidentiality, evaluation results shall be made publicly available. The means of communication shall be readily identifiable and accessible. *(See chapter 5)*
13. The evaluation results shall be transmitted in an appropriate form to meet the needs of decision-makers, beneficiaries and citizens and to fulfil the obligations for reporting in accordance with the legal bases. *(See chapter 5)*

C. Evaluation process

1. Each evaluation shall be devised as a project comprising three separate stages: design, implementation and exploitation. The project manager shall establish who is to take part in these three stages in accordance, where appropriate, with the internal rules and procedures applicable. *(See chapter 3)*
2. The points of importance to the different departments concerned shall be examined when the questions for the evaluations are established. *(See section 3.2)*
3. When designing any (internal or external) evaluation project, the purpose of the evaluation must be clearly and accurately defined. The evaluation project shall comprise the following points: *(See chapter 3)*
 - a. the background to, reasons for and aims of the evaluation,
 - b. for whom it is intended and who will use it,
 - c. the scope of the area evaluated,
 - d. the key questions,
 - e. details of the information available,
 - f. the reports,
 - g. the deadlines.
4. In the event of external evaluation, the evaluation project must also include the contractual, financial and administrative clauses and quality criteria. Specifications shall formalise above points. *(See chapter 3)*
5. The evaluation function shall decide on the advisability of setting up a steering group for each evaluation. *(See chapter 4)*
6. The steering group shall be set up and perform its tasks in compliance with the rules on conflicts of interest and the directives on public contracts. *(See chapter 4)*
7. The tasks of the steering group shall be, as a minimum, to: *(See chapter 4)*
 - a. facilitate the evaluator's access to the information needed to perform his/her work;
 - b. support the evaluation work, particularly from the methodological standpoint;
 - c. take part in assessing the quality of the evaluation at the appropriate juncture.

8. Evaluation processes shall conform to these standards which are designed to ensure their quality.
9. The evaluation project manager shall be responsible for the conduct of the evaluation.
10. The evaluation project manager shall systematically assess the quality of the evaluation on the basis of the quality criteria laid down in the specifications.
11. The persons carrying out evaluation activities shall be free of all constraints which could jeopardise their objectiveness and shall behave honourably in their relations with all stakeholders.
12. The evaluator's independence in his/her work must be respected and the evaluation results must not be interfered with.

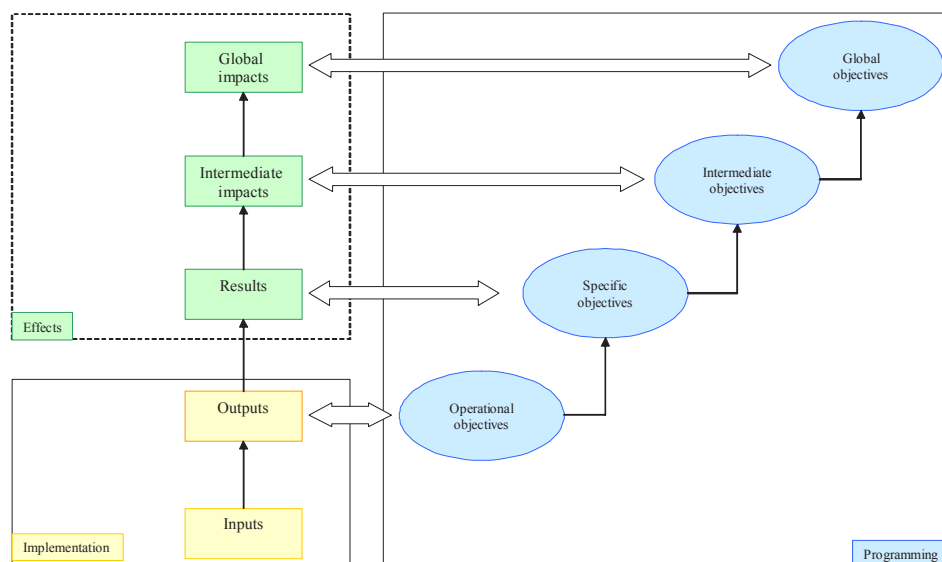
D. Quality of reports

1. The substance of the evaluation reports shall be relevant, based on rigorous analysis, meet the quality criteria laid down in the specifications and comply with the deadlines. *(See section 4.2.3)*
2. The evaluation reports shall describe the purpose of the evaluation and its context and also the objectives, questions, procedures, results and reasoned conclusions of the evaluation, so as to make available the essential information in an easily understandable form. *(See sections 4.2.3 and 5.1)*
3. The report shall describe the information sources in such detail that the correctness of the information can be assessed. The data collected or selected shall be adapted to the methodologies used and be sufficiently reliable for the expected use. *(See section 4.2.3)*
4. The prospects and reasoning on which interpretation of the results is based shall be described and explained. The results should follow on logically and be substantiated by data analysis and interpretations based on carefully-presented explanatory hypotheses. *(See section 4.2.3)*
5. The final evaluation report shall present the results and conclusions of the evaluator and the tenor thereof shall not be amended without his/her agreement. *(See section 4.2.3)*
6. The conclusions and any recommendations shall be rigorous and not distorted by personal or partisan considerations. The recommendations shall be comprehensible, useful, applicable and detailed enough to be brought into effect. *(See section 4.2.3)*

ANNEX B: DIFFERENT ELEMENTS AND CONCEPTS OF INTERVENTION LOGIC

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. Given that the main focus of evaluation is on the effects (i.e. results and impacts) of an intervention, the most relevant objectives by which performance is assessed are those at the specific (especially mid-term evaluations), intermediate and global level:

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 (or *intermediate impacts*) 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.

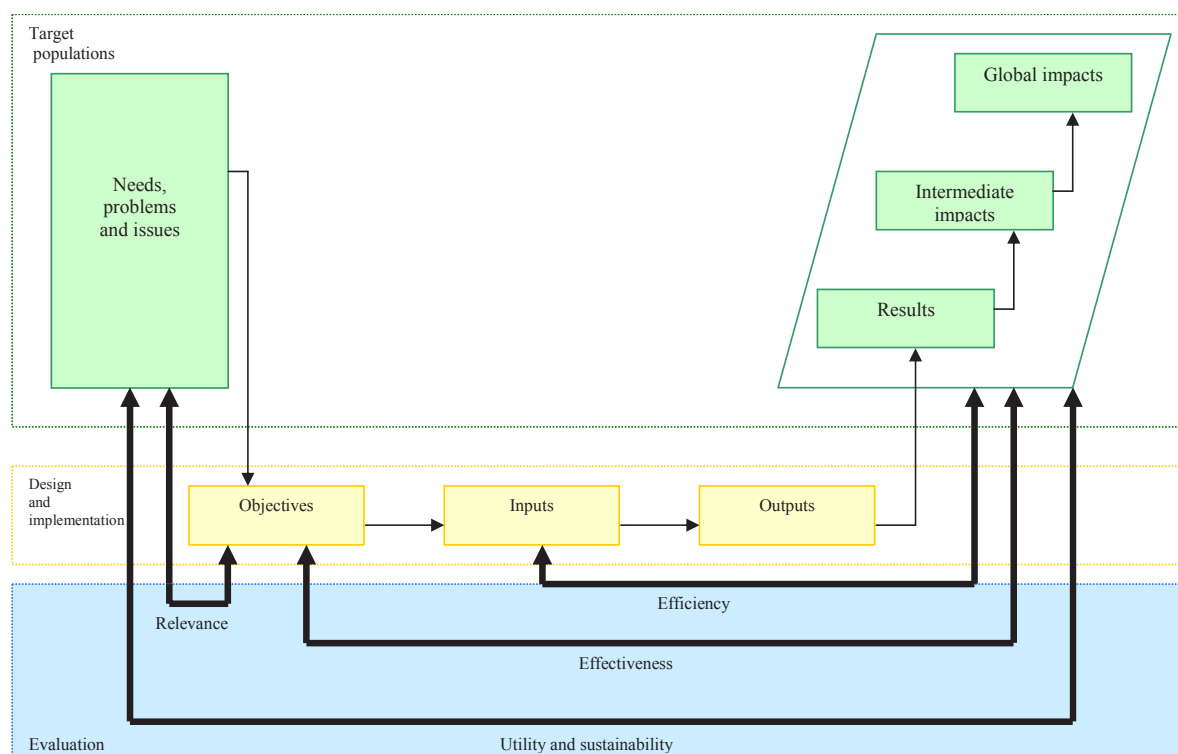
When choosing objectives, following five criteria can be used: Systematic; Measurable; Accepted; Realistic; Time-Dependent.

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) and costs for third parties (Member States, intermediary organisations).

Outputs: Output is defined as a product, which is delivered by the Commission services. 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.

Results/intermediate impacts: Immediate or initial effect / outcome of an intervention.

Global impacts: Longer-term effects / outcomes of an intervention.



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.

Asking the right questions when addressing the evaluation concepts:

To what extent is an intervention relevant in respect to needs, problems and issues identified in target groups? Some time after the initial implementation, the rationale that initially gave rise to the public intervention has to be verified to assess if the strategy remains relevant given the possible evolution of the situation (i.e. evolving needs, problems and issues).

To what extent do the effects induced by an intervention correspond with its objectives as they are outlined in the intervention strategy? A major element in judging the success of an intervention is to assess its effectiveness in terms of the progress made towards the attainment of pre-determined objectives.

How economically have the resources used been converted into effects? In addition to ascertaining if an intervention has attained its objectives, it must also be assessed on the basis of how much it cost to attain them. Hence an assessment of the *efficiency* of intervention is required.

How do the effects of an intervention compare with the wider needs of the target populations/geographical zones? Over and above those effects that correspond with the stated objectives of an intervention, other effects may occur that may be either negative or positive (i.e. unplanned or unexpected effects). An assessment of these provides the basis of a broader assessment of performance on the basis of an intervention's *utility*.

To what extent can any positive changes resulting from the intervention be expected to last after it has been terminated or when beneficiaries are no longer supported? While some interventions merely support certain activities that would otherwise not occur, others may be designed to bring about lasting changes within a target public, geographical zone, etc. An assessment of the latter provides the basis of the *sustainability* of an intervention's effects.

ANNEX C: DIFFERENT VARIANTS OF EVALUATION

(Evaluation at different stages of an intervention)

Ex ante evaluation as an aid to planning and programming

Ex ante evaluation is essentially an aid to planning and programming. Two key advantages result from this process:

- it allows a proper appreciation of whether the proposed level of funding and resources are in accordance with the expected results and impacts, and
- reliable ex post evaluation and hence accountability for results and impacts, is largely dependent on the quality of the preparation of the intervention at the outset.

It is often conducted by operational units with methodological assistance from the evaluation function and possibly from outside consultants.

For ex ante evaluation to be carried out effectively it should be done pragmatically and be proportionate in scale to the intervention that it supports. Usually implemented under tight time constraints imposed by the planning process, it enables a structured reflection on the following questions:

What is the rationale for public intervention? An analysis of needs, problems and issues, both short and long term, to be addressed of the main target groups will provide the response to this question.

What are the effects that the intervention intends to produce and are they relevant? The intended effects of an intervention correspond with its objectives, and for the latter to be relevant they should address the needs problems and issues identified in the above step.

How is the intervention intended to lead to the attainment of the above objectives and how will progress be assessed in this respect? A process of making explicit the hypotheses about how an intervention is expected to attain its higher level objectives allows a hierarchy of objectives to be elaborated and relevant indicators to be identified.

What is the most appropriate delivery mechanism to use? There may be a number of different means to attain objectives and these should be assessed for their suitability in this regard. This assessment will appreciate for each delivery option the likely risks that may result in the non-attainment of objectives.

What is the value added of Community intervention? This step involves an assessment of the extent to which Community involvement is likely to add value to interventions of other administrations, organisations, institutions, etc. (i.e. in being complementary to and coherent with them).

To what extent is the intervention likely to be cost-effective? An assessment of the cost of the proposed intervention in relation to the objectives to be attained will facilitate reflections on the attainability of objectives for a given level of resources, the suitability of different delivery mechanisms, etc.

How will monitoring and evaluation of the intervention be organised and implemented? Arrangements for monitoring the implementation of an intervention through the production of

quantitative data on resources used, outputs delivered and results on beneficiaries need to be planned at the outset. Similarly, given the long lead times needed to prepare, contract and execute evaluations, a clear evaluation plan covering timing, focus, responsibilities and delivery is a necessity

A review of evaluation reports and other appropriate sources (studies, audit reports, etc.) is also an essential task of ex ante evaluation which provides an opportunity to learn lessons on the basis of past experience.

An ex ante evaluation will not necessarily result in a single reporting document. The results of the process should however be presented in an abridged form in the explanatory memorandum and in the legislative financial statement accompanying the proposal.

Interim and ex post evaluation for learning and accountability

Interim and ex post evaluations, as the names suggest, take place respectively during and after the implementation of an intervention. The results of these evaluations are used to support management, organisational learning, planning and policy making and for accountability purposes. Ex post evaluations in particular, but also interim evaluations, tend to be conducted by external evaluators in order to render their conclusions more credible for accountability purposes.

The focus of evaluations of this nature is first and foremost on the effects of an intervention on target publics/geographical areas (particularly in the case of programmes, financial and legal instruments with socio-economic or environmental objectives) drawing primarily on both quantitative and qualitative data collected. Furthermore, evaluators analyse implementation arrangements and processes in order to assess how the latter affect the attainment of objectives and to be able to provide concrete and practical recommendations for improvement. At the heart of an evaluation of this nature is an assessment of **causality**, i.e. an analysis of the extent to which observed effects on target publics/geographical areas are likely to be a consequence of the intervention.

Interim and ex post evaluations are structured around a number of key issues that direct the process of data collection and analysis. Firstly, an intervention should be designed to respond to identified problems, needs and/or issues that manifest themselves in geographical area/target populations that are potential recipients of assistance. If the objectives of an intervention address the needs, problems and/or issues identified, the intervention strategy can be judged as respecting the criterion of **relevance**. The implementation of the strategy depends firstly on the mobilisation of sufficient and appropriate resources (or inputs) that are used to supply outputs, and secondly where relevant on the demand for these outputs from addressees. However, the focus of an evaluation is by definition mainly on the effects of an intervention on target areas/populations and whether these effects:

- correspond with objectives (**effectiveness**);
- are achieved at reasonable cost (**efficiency**);
- correspond with needs, problems and issues over and beyond those embodied in stated objectives (**utility**);
- are likely to continue into the future in the absence of assistance (**sustainability**).

Furthermore, the focus of an evaluation differs according to whether it takes place at the intermediate or ex post stage of the programming cycle. The table below presents these variations.

Issue	Interim evaluation	Ex post evaluation (on-going evaluation of a policy)
Relevance	Yes	No
Effectiveness	Yes in respect to implementation and early effects	Yes
Efficiency	Yes in respect to implementation and early effects	Yes
Utility	Sometimes	Yes
Sustainability	No	Yes

The final output of an interim or ex post evaluation is always a formal report (and various derivative products), which is not necessarily the case for ex ante evaluation.

In reality, the three-stage evaluation cycle has to be adapted to the specific circumstances of an intervention, and hence a single evaluation may be conducted at a critical moment in the implementation of an intervention that incorporates elements of ex ante, interim and ex post evaluation exercises, assessing past performance and examining issues of relevance to its future design.

ANNEX D: EVALUATION OF AGENCIES

This Annex contains an indicative list of generic evaluation questions for the evaluation of Community Agencies (Section 1) as well as an overview of the existing legal provisions concerning such evaluations (Section 2).

1. Indicative list of generic evaluation questions for the evaluation of Agencies

This Section contains a set of generic evaluation questions (and associated success criteria) that are meant to address key issues that are of interest to one or more of the following stakeholders: the Agency's management; the Commission (as the body having proposed the setting up of the Agency and expecting its contribution to the delivery of a given Community policy); the European Parliament (as the Agencies' discharge authority), etc.

Which of the questions contained in sub-section 1.2 will actually be retained in the terms of reference for an interim or ex-post evaluation of an Agency will not only depend on the information needs of its key stakeholders, but also on the *commissioner* of the evaluation, i.e. either the Agency itself or its *DG de tutelle*. It is also conceivable that *two separate* evaluation exercises are carried out: one launched by the Agency with a view to assessing the degree of attainment of its set objectives, and identifying ways of improving the efficiency and effectiveness of its operations; and one launched by the Commission, with a specific focus on the actual value added that the Agency provides as a means of implementing Community policy, and on coherence of the Agency's objectives/activities with regard to other relevant objectives/activities at Community/Commission level (including those at a more strategic level).

In any event, every effort should be made to ensure adequate co-operation and co-ordination between the Agency and its DG de tutelle, so as to maximise the synergies of any evaluations launched by either of them.

1.1. Methodological comments

Given the considerable divergence in the nature and tasks of Agencies, the below generic evaluation questions:

- are phrased in very general terms;
- are not meant to be exhaustive;
- may not be relevant in their entirety for each Agency and/or each individual evaluation exercise.

Hence, in the preparation of future Agency evaluations of an intermediary or ex-post nature, these questions will have to be carefully adapted to the specific characteristics of individual agencies as well as to their specific information needs.

The recommended approach towards making these questions more specific, focused and relevant to a particular evaluation exercise consists of:

- using the generic evaluation questions Q1 to Q9 – which describe the issues at stake in rather general terms – as a starting point (while adapting them to the specific situation);
- developing a set of success criteria in order to describe the expected result or impact of particular activities/measures/factors that appear to be relevant in the context of the overarching generic evaluation question (some indicative examples are already given in sub-section 1.2 below⁶²);
- defining quantitative or qualitative indicators for each success criterion, so as to be able to measure to what extent the expected output or impact has been achieved;
- defining target levels and/or baselines for these indicators, so as to be able to assess whether a certain activity/measure has been successful or not.

At the stage of drafting the Terms of Reference, the evaluation officer may realistically be expected to define the evaluation questions and to develop a first set of success criteria. In the tender documents and in the inception phase of the evaluation, the evaluators should be asked to validate these questions/criteria and to make proposals for further criteria as well as for indicators and target levels/baselines. All of this will then have to be approved by the manager(s) of the evaluation during the inception phase.

Evaluators should also be asked to analyse and describe the Agency's key inputs, outputs, results and impacts as well as the conceptual relations between these elements (i.e. the Agency's intervention logic).

1.2. Evaluation Questions

Relevance

Q1: To what extent are the objectives of the Agency in line with the needs of the addressees/beneficiaries of its activities and/or the social, economic or environmental problems the Agency is meant to address?⁶³

Added value

Q2: To what extent does outsourcing to the Agency provide added value compared to possible alternative options of implementing the Community policy in question (e.g. through the Commission Services themselves, an Executive Agency, the contracting out of individual tasks, etc.).⁶⁴

Indicative list of success criteria:

- a. Specialised expertise or know-how acquired;

⁶² Some of these criteria may be so important for a particular Agency that it may be decided to (re-) transform them into fully-fledged, self-standing evaluation questions.

⁶³ For guidance on how to carry out problem analysis and needs assessment, see the Commission's practical guide on ex-ante evaluation:

http://europa.eu.int/comm/budget/evaluation/Key_documents/evalguides_en.htm

⁶⁴ The cost-benefit analysis or ex-ante evaluation carried out at the time of creation of the Agency may serve as a baseline for this evaluation question. Due account should be taken of changes that have occurred since then, with regard to both context and the Agency itself (e.g., mandate, tasks, etc.).

- b. Timely and high-quality response given to ad-hoc information/service requests made by EU institutions or other stakeholders;
- c. Credibility of the Agency's activities/outputs enhanced as a result of greater independence;
- d. Additional and/or more effective ways of stakeholder involvement (e.g. Member States, Social Partners, etc) ensured;
- e. Comprehensive networks for the gathering and exchange of information and best practice established;
- f. Favourable impact on the budgetary framework of the European Union achieved;
- g. Flexibility in the implementation of outsourced tasks achieved;
- h. Proximity to final beneficiaries/addressees ensured.

Coherence, complementarity, synergy

Q3: To what extent are the elements of the Agency's intervention logic complementary, mutually supportive and non-contradictory? To what extent do the objectives and activities of the Agency support or contradict those of other public interventions?

Indicative list of success criteria:

- a. Coherence ensured with regard to the overall strategic objectives at Commission and Community level;⁶⁵
- b. Consistency ensured with regard to the Commission's and Community's key objectives in the Community policy area to which the Agency pertains;⁶⁶
- c. Internally coherent hierarchy of objectives established (from Constituent Act down to Annual Work Programme);
- d. Consistency ensured between high-level objectives in Constituent Act and the resources, responsibilities and competences entrusted to the Agency;
- e. The objectives of the Agency complement and/or support those of other public interventions, in particular those of [relevant Community Agencies], [the European Commission], [other EU institutions], or [relevant national or international organisations/bodies active in the specific policy area];⁶⁷
- f. The activities of the Agency complement and/or support those deployed by other public interventions, in particular those of [relevant Community Agencies], [the European Commission], [other EU institutions], or [relevant national or international organisations/bodies active in the specific policy area];⁶⁸

⁶⁵ For instance, strategic objectives defined in the context of the Commission's Strategic Planning and Programming (SPP) cycle.

Specify policy area, key legislation, policy and operational documents.

⁶⁷ In order to focus this criterion as much as possible, the most relevant Agencies, European Commission policies/legislation/documents, Council or European Parliament documents, as well as international organisations/bodies have to be carefully selected and specified.

⁶⁸ Ibid.

- g. Effective mechanisms are in place in order to ensure co-ordination/co-operation with the [Commission Services], [other EU institutions], [Community Agencies], [International Organisations/Bodies] most concerned.⁶⁹

Efficiency and effectiveness⁷⁰

[Effectiveness of Agency-specific activities]

Note: Evaluation questions relating to effectiveness are intimately linked to the specific objectives of Agencies. Detailed questions will hence have to be established on a case-by-case basis. For the same reason, no success criteria are suggested for these questions. Evaluators should be asked to identify not only the results of the activities/interventions in question but, wherever possible, also their impacts (including longer-term ones).

Q4: To what extent has the Agency achieved the objectives set out in its work programme(s)? In particular:

A. *[specific question relating to attainment of key objective A]*

B. *[specific question relating to attainment of key objective B]*

C. *[specific question relating to attainment of key objective C]*

X. To what extent has the Agency succeeded in setting up its Network(s) [specify] in accordance with its mandate? What has been the contribution of this/these Network(s) towards the attainment of the Agency's objectives?

[Questions pertaining to the efficiency of Agencies]

Note: Again, specific questions will have to be established on a case-by-case basis, depending on the special characteristics of Agencies and their information needs. In any event, such questions will have to link and compare key outputs/results of Agencies with the inputs needed to produce them (financial and human resources, legal and administrative framework, etc.).⁷¹

Q5: To what extent has the Agency achieved ... [specify output/result] at a reasonable cost in terms of financial and human resources deployed? In particular:

A. *[specific question A]*

B. *[specific question B]*

C. *[specific question C]*

X. To what extent do the Agency's procedures in the areas of financial and human resources management affect the efficiency of its operations?

⁶⁹ Specify the most relevant services/bodies.

⁷⁰ Over and above the questions and success criteria identified under this heading, evaluators should be asked to identify any other key factor that is conducive to, or hampers, the efficiency or effectiveness of the Agency's operations.

⁷¹ The ratio between the Agency's administrative and operational expenditure may be an indicator in this context, to be compared to an appropriate benchmark (e.g. other Agencies) or to a baseline such as the previous performance of the Agency (e.g. past 5 years) or the expectations at the outset.

[Organisational set-up and decision making]

Q6: To what extent does the Agency's organisational set-up contribute to the effectiveness and efficiency of its operations?

Indicative list of success criteria:

- a. The size, composition and context of the Board strike a reasonable balance between the need to retain an effective decision-making body and the need to ensure stakeholder representation, involvement and commitment;
- b. Board provides clear strategic direction and sets priorities;⁷²
- c. *[Depending on information needs and Agency-specific characteristics, criteria concerning other components of the organisational set-up (e.g. scientific committees, etc) could be added.]*

[Internal management systems]

Q7: To what extent do the Agency's management systems and processes contribute to the effectiveness and efficiency of its operations?

Indicative list of success criteria:

- a. Specific, realistic and operational objectives as well as indicators for outputs, results and impacts contained in work programme(s);
- b. Monitoring system allows the Agency to collect relevant data on inputs, outputs, results and impacts (where possible);
- c. Monitoring data and evaluation findings fed back into decision making;
- d. Activities (and resources) of Agency focused on priority objectives;⁷³
- e. External communication and dissemination strategy established;
- f. Internal communication contributes to ... *[specify key objective(s) of internal communication policy]*;
- g. *[Depending on specific information needs, additional criteria should be added to capture other key management systems and processes likely to affect the efficiency and effectiveness of Agencies.]*

Utility

Q8: To what extent do the results and impacts of the Agency's activities correspond to the needs of its clients/beneficiaries/addressees and/or the social, economic and environmental problems it is meant to address?

⁷² Examples of qualitative indicators: Delegation of less strategic tasks to an Executive Bureau (or similar structure); Board meetings are focused on issues of strategic importance; etc.

⁷³ This criterion relates to the "system" that ensures that activities are prioritised and resources allocated accordingly. For concrete activities, the match between objectives and resources could be looked at on a case-by-case basis under Q4 (effectiveness), since the availability (or lack of) resources may be an explanatory factor when assessing the extent to which set objectives have been reached.

Indicative list of success criteria:

- a. The Agency's activities have made an actual contribution to the delivery of Community policy in the area to which the Agency's activities pertain;
- b. The clients, beneficiaries or addressees of the Agency are satisfied with its activities and/or outputs.
- c. The Commission has made full use of the Agency's potential to contribute to the delivery of Community policy.

Q9: To what extent have the activities of the Agency resulted in any unintended/unplanned results and impacts (both desirable and undesirable)?

Note: Unintended/unplanned effects may concern inter alia the target population, society at large, the EU institutions, or result from the local setting of the Agency. For instance, outsourcing to the Agency may have resulted in reduced visibility of the Community as promoter of the programme/policy concerned. It may also have led to the situation that the Commission has not been able to acquire and/or maintain an adequate level of know-how inside the institution. The choice of the geographical location of the Agency may have had manifold implications (e.g. on costs, co-funding and other support by the host state, recruitment, visibility, image, etc.).

No "success" criteria are proposed here because the effects in question have not been intended.

2. Legal provisions concerning the evaluation of Agencies

a) Agencies covered by Article 185 of the Financial Regulation

Most of the existing Community Agencies are subject to the "framework Financial Regulation for the bodies referred to in Article 185 of the Financial Regulation".^{74 75} Article 27(4) of that Regulation obliges such Community bodies to "*regularly carry out ex ante and ex post evaluations of programmes or activities. Such evaluations shall be applied to all programmes or activities which entail significant spending, and evaluation results shall be sent to the management board*".

In other words, the general rule applicable to such Agencies mirrors exactly the rule that is applicable to the Commission's own Services under Article 27(4) of the general Financial Regulation⁷⁶. However, the obligation to evaluate lies on the Agency *itself*.

It should be noted that, while the *general* Financial Regulation contains, in Article 21 of its Implementing Rules⁷⁷, also more detailed provisions on evaluation (in particular concerning the content of ex ante evaluations and the timing of interim and ex post evaluations), no such provisions are contained in the framework Financial Regulation for Article 185 bodies. [Agencies are free to adopt their own implementing rules.](#)

⁷⁴ Commission Regulation N° 2343/2002 of 23 December 2002.

⁷⁵ Under the first pillar of the EU Treaty, the Agencies concerned are CEDEFOP, EUROFOUND, EEA, ETF, EMCDDA, EMEA, EU-OSHA, CdT, EUMC, EAR, EFSA, EMSA, and EASA. Excluded are OHIM and the CPVO, both of which do not receive subsidies from the Community budget. As regards the third pillar of the EU Treaty, EUROJUST is covered by the Article 185 framework Financial Regulation as well.

⁷⁶ Council Regulation N° 1605/2002 of 25 June 2002.

⁷⁷ Commission Regulation N° 2342/2002 of 23 December 2002.

Similarly, Agencies enjoy a considerable degree of autonomy with regard to the adoption of an [Internal Control Standard on evaluation](#) (comparable to the Commission's ICS N°23 on audit and evaluation). According to Article 38(4) of Commission Regulation N°2343/2002, *"the authorising officer shall put in place, [...] on the basis of equivalent standards laid down by the Commission for its own departments, and having due regard to the risks associated with the management environment and the nature of the action financed, the organisational structure and the internal management and control systems and procedures suited to the performance of his/her duties, including where appropriate ex post verifications."*

ICS n°23 therefore does not apply *directly* to Agencies, who have a certain room for manoeuvre in order to take account of the specific conditions under which they operate. However, the Commission's rule still serve as a *reference* (i.e. have to be taken "as a basis") for this exercise. In this context, it should be noted that Agencies are now under the umbrella of the Commission's Internal Audit Service (IAS) with regard to internal audits, with the IAS expecting Agencies to comply with international best practices in matters of internal control.

As far as traditional agencies are concerned, the [role and responsibility of the Agencies' "DG de tutelle"](#) within the Commission Services is rather limited (in contrast to executive agencies, see point d below): The responsibilities of the Commission are limited to those that are laid down *explicitly* in the Constituent Acts of the Agency concerned. Most of these Acts specify that it is the *Agency's* responsibility to conduct/commission evaluations, with the Commission's role being limited to the provision of information or agreement to the terms of reference proposed by the Agency. That having said, DGs are nevertheless required to take account of "their" Agencies' operations when evaluating the ABB activity under which these Agencies are financed.

b) [Traditional Agencies not covered by Article 185 of the Financial Regulation](#)

The two self-financed Agencies, OHIM (Office for Harmonisation of the Internal Market) and CPVO (Community Plant Variety Office), are not covered by Article 185 of the Financial Regulation, and adopt their Financial Regulations independently. These Agencies are not covered by the Commission's rules on evaluation, given that Article 27 of the general Financial Regulation only concerns those institutions that are defined in Article 1 of that Regulation. OHIM and CPVO would only be obliged to evaluate their activities if their respective Constituent Acts required them to do so, which is however not the case.

c) [European Regulatory Agencies](#)

According to the proposed framework for European Regulatory Agencies⁷⁸, any newly created regulatory agency would also be covered by the Article 185 framework Financial Regulation (to the extent that these agencies receive subsidies from the Community budget). Hence, the evaluation requirements described under point (a) would apply.

⁷⁸ Communication on the operating framework for the European Regulatory Agencies, COM(2002)718.

d) Executive Agencies

The statute for Executive Agencies⁷⁹ provides, in its Article 25, that the *Commission* has to ensure external evaluations at three-yearly intervals (which may even result in the decision to wind up agencies that are no longer justified with a view to sound financial management).

Over and above this general evaluation requirement (imposed on Commission Services), the "Guidelines for the establishment and operation of executive agencies financed by the general budget of the European Communities"⁸⁰ suggest that Services may entrust the agencies with certain tasks in the field of evaluation: For instance, executive agencies can be made responsible for evaluations on *projects* managed by them, and they could/should also play a certain role in the (preparation of the) evaluation of the *programmes* in the context of which they have been established.

The guidelines also indicate that executive agencies have to respect the Commission's Internal Control Standards, however, only to the extent that these standards are "relevant" for the agency in question (i.e. adapted to the agency's management environment). In all events, "*the directors must put in place management and internal control systems adapted to the performance of tasks entrusted to the executive agency in such a way as to ensure the legality, regularity and sound financial management of the operations it performs.*"

⁷⁹ Council Regulation N° 58/2003 of 19 December 2002.

⁸⁰ See <http://www.cc.cec/budg/lex/en/others/execution/executive.htm>

ANNEX E: EVALUATION TOOLS AND TECHNIQUES

This annex presents a number of tools and techniques often found in evaluation methods employed in the context of the Commission's interventions.

Tools and techniques adapted to the structuring phase of an evaluation

The structuring phase of an evaluation covers the period from the drawing up of the terms of reference by officials to the delivery of the inception report by the evaluator.

- **Intervention logic analysis**

Intervention logic can be defined as a set of hypothetical cause and effect linkages that describe how an intervention is *expected* to attain its global objectives. At the design phase of an intervention, making explicit the logic of a planned intervention allows the construction of a hierarchy of objectives and the identification of key indicators for monitoring and evaluation (See Guidelines for ex ante evaluation). When conducted at a later stage, an analysis of this kind can be used to support the drawing up questions for intermediate and ex post evaluation. If an evaluation is narrowly focused (for example on a small number of instruments) or the intervention is of limited complexity, the **logical framework** technique can be used to conduct this type of analysis. In other cases more elaborate **logic models** (see section 3.2 and the **example from DG AGRI**) have to be constructed that are able to capture in a qualitative fashion the complexity of an intervention's logic. Data sources for constructing logic models are key documents, in particular the legal base of an intervention and supporting policy documents and key informants, a number of which will usually be members of the steering group.

- **SWOT analysis**

SWOT analysis is a diagnostic tool that can be used in ex ante evaluation to ensure that an appropriate intervention strategy is developed on the basis of a structured, comprehensive and transparent analysis. A SWOT analysis examines the *strengths* and *weaknesses* of, for example, a region, a sector, an organisation and assesses the future *opportunities* and *threats* to it. This type of analysis is most likely to draw on documentary (e.g. previous evaluations, studies, etc.), statistical sources of data and interviews with key informants.

Tools and techniques for data collection

- **Questionnaire surveys**

Questionnaire surveys are best adapted to collecting facts and opinions from beneficiaries during intermediate and ex post evaluations. They can be most effectively implemented when an up-to-date list of beneficiaries and their contact addresses is available. Questionnaires are most effective when made up of closed-end questions that give rise to a choice of quantitative or qualitative responses that can then be analysed to produce quantitative data about an intervention's effects, delivery mechanisms, etc. The choice of questions and the need to be able to predict the most relevant responses to them requires a good level of prior knowledge of the target groups, the intervention, its effects, etc. or some thorough preparatory work to acquire this knowledge (e.g. through interviews or focus groups – see below). A survey questionnaire can be designed to generate bottom-up quantitative data about the counterfactual situation, i.e. what addressees would have done if support had not been provided through the intervention. While results from a sample can be safely generalised to the population when more than 1000 observations are analysed, an analysis based on as few as 30 observations can provide evaluators with usable information. Conducting a successful questionnaire survey

requires a good knowledge of questionnaire design and sampling techniques, as well as an understanding of the practical issues relating to the administering of a questionnaire (e.g. by letter, fax, email, webpage, etc).

- **Interviews**

Interviews with key informants are suitable for all evaluation situations, being adapted to collecting qualitative data of both a retrospective and prospective nature. They may be structured in character, taking a form similar to that of a questionnaire, semi-structured with the interview focusing on providing detailed information on a number of key themes whilst leaving the possibility to explore sub-themes and to develop new themes, or unstructured so as to allow areas of interest to be developed and explored in the course of the interview. Interviews are especially useful for gathering in-depth information about process-related issues, i.e. how events occur and why. In addition to being a standalone tool for collecting data in the field they can also be used for preparing other tools and techniques for implementation. For example, documentary research in the structuring phase of an evaluation can be complemented by interviews with key informants, who are able to provide a strategic overview of an intervention, detailed input into the analysis of intervention logic, etc. An interview programme should be designed to collect data from the range of different viewpoints on a particular subject.

- **Focus groups**

This survey technique is based on a small group discussion. The particularity of this tool is its ability to harness the dynamics of the people, who are affected by an intervention in order to generate rich, in-depth information on a specific subject. Focus groups can be used as a means to collect both retrospective and prospective data of a qualitative nature, for example about existing or past processes and effects, or about possible future events and situations that may arise as a consequence of a yet to be implemented intervention. They can also be used to prepare other data collection tools, such as questionnaire surveys, as well as to provide a forum in which previously collected data can be analysed and interpreted collectively by the participants. Focus groups should be representative in the sense that participants bring the variety of different viewpoints to bear on a particular subject. Groups should be homogenous in the sense that their composition is conducive to frank and meaningful exchanges.

- **Case studies**

Case studies can be considered as in-depth, self-contained studies embedded within a larger study or evaluation. Case studies are usually used to generate and analyse data of a retrospective nature about a particular entity (the case-study object), for example, a region, a firm, an institution, etc. There are three main types of case studies: exploratory, descriptive and explanatory. In the first instance, a case study or a number of case studies are used to develop a better understanding of key issues, so as to prepare a subsequent more focused round of data collection and analysis. Descriptive case studies are used to illustrate less in-depth analyses with specific and detailed information about the case study objects in question. Explanatory case studies are used to test hypotheses across a relatively limited number of cases, searching to replicate specific results or processes. Case studies will collect and analyse both quantitative and qualitative data about the context in which a case exists as well as the case itself, using a diversity of information sources and tools with a view to searching for patterns across data, and with the aim of triangulating (i.e. use of data collected using different tools and from different sources, and/or analysis from different theoretical perspectives and by different analysts) to underpin the validity of findings. The aim of a case study programme (e.g. to illustrate good practice, to understand what happens in typical cases, to explain why a certain type of case is a success...) determines the choice of cases to be studied, rather than the need to establish a “sample” that is representative of the population under study.

Tools and techniques for data analysis

• **General tools and techniques for the qualitative analysis of data**

Qualitative analyses are most effective when they are undertaken in a structured and transparent way, respecting the principle of triangulation. At the level of data collected by individual data collection tools and techniques, qualitative analytical approaches can be used to analyse both purely qualitative data generated through interviews, focus groups, documentary reviews, etc. and mixed quantitative and qualitative data, as in the case of case studies. Such analyses aim to produce plausible explanations of phenomena that are solidly “grounded” in the data collected.

Given that most evaluations of EU interventions are based on the collection and analysis of both quantitative and qualitative data covering implementation processes, delivery mechanisms, addressee behaviours and other effects, the cross-analysis of data is by necessity usually qualitative in nature, consisting of searching for and matching patterns across the data produced during the initial analyses.

A number of approaches exist that are used for sorting and grouping data so as to facilitate an exhaustive search for patterns within it:

- **Coding and abstraction.** The identification of categories of concepts that are used to label data (coding), the grouping of linked categories of data and the conceptualisation of the latter at a higher level of abstraction to produce conclusions.
- **Data matrices.** The identification of key themes or dimensions and the sorting of data in respect to them, hence making patterns across data easier to draw out.
- **Frequency counts.** The identification of key themes and assertions and counting the number of times that they occur in the data.
- **Time-series qualitative data analysis.** The chronological ordering of data to provide an account of activities and events in such a way as to identify causal relationships.

The overall philosophy of qualitative analysis can be likened to that of patient detective work that aims to uncover corroborating facts and figures and the meticulous preparation of evidence-based argumentation that would be subsequently presented in a court of law.

• **General tools and techniques for the analysis of quantitative data**

Quantitative data generated through survey questionnaires, by monitoring systems and statistical agencies can be analysed using a variety of techniques that go further than the production of **descriptive statistics** (frequencies, mean values, etc), for example:

- **Bivariate and multivariate analyses** are used as a means of establishing if there is an association, or correlation, between one variable and another (or others), i.e. if a change in one variable can be shown to be linked to a change in other variables, either through cross-tabulation or regression analysis.
- **Factor analysis** is often used to establish groups of entities (individuals, firms...) on the basis of similarities between them or between variables. It is therefore particularly useful for constructing typologies, classes within a population, etc. It can also be used to explore information contained in large multivariate data sets and to reduce its complexity.

- **Variance analysis** across groups of observations (for example from a “control” and a “treatment” group in the case of comparison groups, see above) is used to assess whether differences between the groups are significant or not.

These techniques are not however particularly useful for the cross-analysis of data in the context of an evaluation since this usually consists of processing heterogeneous data that is both quantitative and qualitative in nature. In this situation a more qualitative approach to analysis is required.

- **Comparison groups**

The basic principle of comparison group approaches is to compare the evolution of a key variable (e.g. income level, company turnover...) in both the addressees of an intervention (the “treatment” group) and the group unaffected by it (the “control” group) usually within the framework of an ex post evaluation. In doing so they can provide a quantitative estimation of the counterfactual situation and hence of net effects. However, a number of practical problems exist that make establishing a control group problematic, in particular: obtaining contact details of non-addressees; persuading them reveal the required information; finding non-addressees that are of a sufficiently similar profile to addressees, and finding non-addressees not indirectly affected by an intervention. A pragmatic approach that can be used when the intervention has relatively minor penetration rate in respect to the total population is to compare actual addressees to the total population of potential addressees (i.e. the “control” group) in an area, sector, etc. In this instance data on the population would be assembled from statistical sources and compared to data on addressees collected through questionnaires. There are number of ways of using a comparison group approaches to analyse the effects of interventions on addressees: before/after trend comparison (considered the most reliable approach, it compares data on a number of occasions over time); before/after comparison (comparing baseline data to data from a single subsequent moment in time); afterwards only comparison (the least reliable approach as it disregards the differences between entities at the outset of the intervention).

- **Econometric and input-output models**

The construction of analytical models requires a high level of expertise and a considerable amount of time. Econometric and input-output (sectoral) models are usually used for prospective analyses, estimating the value of key effects of an intervention over some period into the future. These tools can be used to create different scenarios, in particular one in which an intervention takes place and one in which the intervention is absent (the counterfactual situation) and hence can be used to estimate an intervention’s likely net effects, in both ex ante and an ex post evaluation situations.

The capacity of a model to produce meaningful estimates of future situations depends firstly on the accuracy of the assumed causal relationships on which the model is built, which by necessity represent a simplified version of a complex reality, and secondly on the stability of these relationships over time.

It is often the case that insufficient time and money are available to develop specific models for evaluations, so existing models when available are sometimes adapted to particular needs. However, one practical way of dealing with time constraints is to undertake a model-based study prior a planned evaluation. The results of the study would then feed into the subsequent evaluation.

- **Expert panels**

Panels of experts can be mobilised to analyse data, in both prospective and retrospective contexts. They are particularly useful for cross-analysing unstructured and heterogeneous data. In this situation experts can be asked to focus their expertise on estimating effects and establishing the counterfactual situation. An expert panel can also be used to judge an intervention on the basis of the results of an evaluation, and to conclude on the extent to which it has been a success or not.

The quality of a panel's analysis or judgement depends on the choice of the most appropriate experts in terms of their knowledge and experience in policy field, their willingness to fully participate in the process and their independence vis-à-vis the intervention being evaluated. Given these requirements and the limited availability of experts, the time needed to prepare and run a panel can extend to a good number of months.

Another way of mobilising expertise is through a **DELPHI survey**. Instead of bringing the experts together physically, this technique allows them to be questioned individually (and hence anonymously if desired thereby reducing the risk bias arising from peer pressure within the group) and in an iterative manner through a centrally administered and evolving questionnaire. The results of this process can eventually be quantified.

Tools and techniques for the judging phase of an evaluation

When intermediate and ex post evaluations have been designed on the basis of terms of reference containing questions that focus on the appropriate generic evaluation issues, and success criteria, indicators and target levels have been established, an explicit basis for judging an intervention exists. Where this pre-established structure is absent, the judgement will be more qualitative and less explicit in nature, unless formal tools are used (See the examples of cost-benefit analysis and multi-criteria analysis presented below). Depending on the approach taken, judgement may be delivered by the evaluator (corresponding with a “managerial” approach that emphasises the independence of the evaluator, ensuring the credibility of the evaluation for accountability purposes) or with a strong input from the steering group (i.e. a “pluralist” approach that emphasises the sharing of knowledge and consensus building among stakeholders, usually with the aim of supporting organisational learning).

- **Cost-benefit analysis**

A full cost-benefit analysis is a major undertaking in terms of both time and money. It is usually used to assess prospectively the overall balance of positive and negative effects that are likely to arise over a period of time from a number of policy options. It consists of identifying the potential “winners” and “losers” among stakeholders, assessing in physical and then in monetary terms their gains and losses over time, and then comparing the overall benefits to costs.

A major advantage of using this tool is that it requires a thorough analysis of both the positive and negative implications of policy options for stakeholders. However, whilst it is often seen as providing an objective basis for making rational choices between policy options, the final results are sensitive to methodological decisions on how to place an economic value on non-economic benefits and costs (e.g. societal and environmental effects), the choice of discount rate, and where to draw the boundary of the analysis (i.e. which stakeholders to exclude). Furthermore, it is difficult to make reliable comparisons across different cost-benefit analyses due to these discretionary methodological aspects unless a considerable effort of methodological standardisation is made.

- **Multi-criteria analysis**

Multi-criteria analysis is more usually used within the framework of prospective analyses and can be employed to support judgements in complex situations, where an intervention (or its various components) is assessed in relation to multiple criteria, and this is one respect by which it differs from cost-benefit analysis. However, there are two other major difference between these tools: multi-criteria analysis does not require the monetarisation of non-monetary effects and allows for the weighting of criteria to reflect differing stakeholder preferences.

ANNEX F: EVALUATION OF POLICIES

Policy Areas & Activities -and non-expenditure instruments

Introduction

The Communication on evaluation⁸¹ broadened the scope of the Commission's evaluation policy and centred it on *activities* including legislative activities and *policies*. This development was further pursued in the framework of the Communication on ABM⁸² and the new Financial Regulation⁸³. On this background, the different treatment of spending and non-spending within evaluation has become increasingly irrelevant because the ABM/ABB framework dissolves the separation between different types of resources related to the Commission's interventions meaning that spending programmes and purely regulatory proposals become merged into activities.

The text below discusses some the aspects of the evaluation of policy, especially where differences arise in respect to programme evaluation. However, while the distinctiveness of policy evaluation needs to be acknowledged, it should not be overstated.

Various aspects of policies

Generally speaking a policy can be understood as an overarching strategy or goal(s) addressed by a number of different policy instruments. Policy instruments, on the other hand, can be considered as set of techniques by which public authorities attempt to ensure support and effect or prevent social change. In this sense, there is a strong emphasis on the dynamic and evolving nature of policies, with individual policy instruments being added, retired or redesigned over time. Policy objectives are also subject of similar modifications, as well as changes in their relative importance

Within the Commission the notion of a *policy* may designate various breadth of scope and levels of complexity, ranging from an overall Commission strategy or objective over a Policy Area to an ABB-activity. In this context a policy will normally embrace a range of instruments. The variety of **policy instruments** available to the Commission include, for example, legislation such as regulations or directives and may involve resource commitments for example in the form of operational programmes; they also include Communications, action plans, etc. However, such is the heterogeneous nature of policy instruments that they differ significantly in *the way in which* they bring about results and impacts and the timescales over which these can be expected to occur.

Analytical implications:

By mapping the 'outputs', results and impacts and their mutual interactions, there are almost certainly added layers of complexity and interaction that arise in the case of a policy. This means that a multitude of cause and effect linkages can be tested and hence an important number of evaluation questions can be posed. In order to keep a policy evaluation well focused, it is helpful to outline this complexity already in the preparatory phase of an evaluation. This can help to identify, in a preliminary manner, the most important effects,

⁸¹ SEC (2000)1051

⁸² SEC (2001) 1197/6&7

⁸³ Council regulation (SEC(2002) 1605) and the implementing rules in Commission Regulation (SEC(2002) 2342)

pathways and bottlenecks for the delivery of the policy that will consequently be considered as priority elements to be examined through the evaluation.

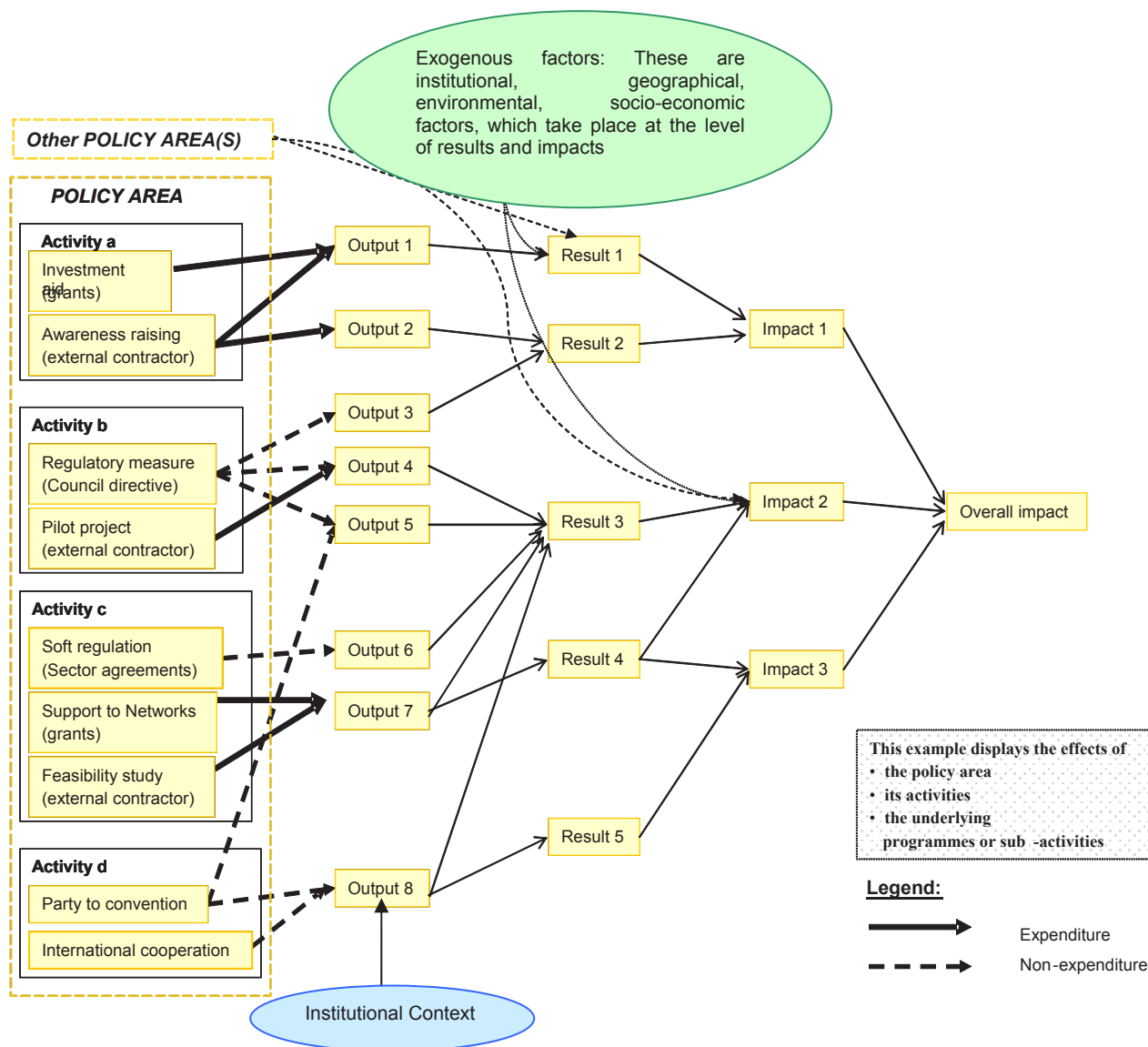
As an example of a way to assess the policies in relation to their higher level objectives consider the following case of a **contribution analysis**. Contribution analysis can be described as a specific analysis undertaken to assess the contribution of an intervention to the impacts it is trying to influence. Its aim is to paint a credible picture about the attribution of an intervention by trying to demonstrate "plausible association", i.e. whether a "reasonable person" on the basis of information gathered and analysed about what has occurred in the framework of an intervention at the input, output and results levels and the context in which the intervention is implemented, agrees that the intervention contributed to effects at the higher levels of impacts. The mapping of intervention logic is a key element of such an analysis.

The following aspects are useful to keep in mind when analysing the effects of policies.

- how a policy instrument is concretely implemented in Member States (N.B. it is in this respect that the concept of an output in the programme sense is less relevant for a policy instrument)
- who is affected (i.e. who are the direct and indirect addressees of the intervention),
- how the intervention is intended to work and how it works in reality (i.e. test the hypotheses on cause and effect represented by the arrows in the diagram below),
- why the intervention produced certain effects and not others (i.e. understanding the factors that facilitate, retard or block the occurrence of effects).

The diagram below illustrates a hypothetical Policy Area. This policy embraces four Activities: one consisting of 2 expenditure programmes, one consisting of 2 non-expenditure measures and furthermore two which comprise a mixture of resource commitments, legislative measures as well as different forms of agreements (soft regulation etc.). This diagram maps **what** the potential effects of the various policy are (cf. the rectangular boxes) and **how** these effects are likely to be related (cf. the arrows). It also demonstrates that as the focus of an evaluation moves from inputs and outputs in the direction of the results and impacts of a truly integrated Activity and/or Policy Area, the more difficult it will be to separate the net effects of each of the constituting policy instruments due to interactions between them and because of the effect of the other measures within the Policy Area and the increasing influence of exogenous factors.

To better understand the diagram below, consider a following example. The Community launches a regulatory measure, such as a Council directive (activity b-1 in the diagram). The directive aims at improving the quality of air (impact) through a reduction in polluting activities (result). The outputs of the intervention could be the implementation of numerous instruments, for example: quantitative limits on pollution levels and fines for breaching them, taxes and levies etc.



In the diagram above it is worth noting that the institutional context intervenes at the level of outputs whereas other exogenous factors (geographical, environmental and socio-economic factors) interact with the effects of the intervention mainly at the level of results and impacts.

This type of mapping will necessarily involve a certain level of simplification in complex cases such as comprehensive Activities or entire Policy Areas. It will for example be a snap-shot of the existing situation at any one moment in time, given the generally evolving nature of policies. Moreover, the mapping will have to focus on those delivery mechanisms explicitly or implicitly foreseen at **the EU-level**, e.g., in the relevant Regulation(s), while for some measures the choice of specific instruments and the way in which they are implemented depends on decisions made at Member State level. It is important to remember that the intervention logic represented in the diagram is, until tested through an evaluation, no more than a series of hypotheses, whatever the type of intervention concerned. It is for this reason that such an analysis is often referred to as *“theory of action”* or *“programme theory”*.

Methodological implications

The diagram and the discussion above suggests evaluation approaches and techniques for expenditure programmes are likely to remain directly relevant for policy evaluation and non-expenditure interventions, even if certain concepts may need reinterpretation or a change in emphasis. This section will comment on four such aspects: firstly the addressees of an intervention, secondly the concept of ‘cost’, thirdly the notion of allocative effects and finally certain analytic elements.

The addressees of the intervention(s)

Any type of intervention should be designed in such a way as to target clearly identified addressees. In the case of a **spending programme**, addressees are usually referred to as beneficiaries as they are intended to *benefit* from the intervention. Moreover, the identity of each individual direct addressee will normally be known to the programme managers because of the need to disburse money to them.

On the other hand, **non-spending interventions** often aim to create advantages and disadvantages (often additional costs) for various addressees, with the aim of reaching one or a number of global (or high-level) objectives. Addressees are therefore not necessarily beneficiaries in this case. The reason for any public intervention should, however, always be to produce net benefits for society (i.e. globally speaking the advantages should outweigh the disadvantages), and an evaluation of a policy will take this issue into account. Some addressees may be known individually, though not in a systematic way as would be the situation in the case of spending programmes, or they may only be known by type to the public authority.

As an example of direct and indirect addressees consider the case of an air quality regulation. This legislation, for example, imposes a tax (a measure in the above diagram) on businesses to promote the use of cleaner technologies. The businesses paying the tax (i.e. the direct addressees) have an incentive to change their behaviour by adapting their activities through the introduction of cleaner technologies (a result in the diagram). Consequently, airborne emissions from production facilities may be reduced (the lower level of impacts) leading to the possible improvement in the air quality for society as a whole (the high level impact). In this example individual members of society are the indirect addressees of the policy.

Mapping the intervention logic therefore is closely linked to identifying both direct and indirect addressees. The above diagram could be usefully complemented by adding information on **who** is affected by the intervention. Such an analysis can be used to gain a better understanding of the possible allocative effects of policy instruments (see below).

The concept of outputs

In a spending programme the output is considered as that which is financed and accomplished with the money allocated to the intervention. In a non-spending or policy context the intervention logic may begin with the initial effect vis-à-vis the addressees. Consider a hypothetical example of a directive which gives the Member States the option between an outright ban of a product or process, or alternatively, various measures to increase security (process, storage, training of operators, etc). The outputs at EU level could in such a case be based on a typology of the key types of such measures in the Member States. On the other hand it would in most evaluation context not be useful to consider the adoption of the directive by Council and Parliament as an output, nor its transposition into law in the Member States. This latter aspect is more a concern of the monitoring of implementation

The cost implications

Another specificity of the evaluation of non-spending measures (i.e. without budgetary costs) is the understanding of the costs to be taken into account in an evaluation. The notion of **efficiency**, whatever the type of intervention, refers to the relationship between the effects induced and the cost to the implementing body (depending on the policy instruments, one or both of the types of cost identified in the column 'Implementing body' in the table below).

On the other hand, addressees may be expressly the target of policy instruments that generate disadvantages and which may therefore lead to them incurring significant costs (of the type indicated in the column "Addressees" below). These costs represent a potentially major category of **negative impact** of a policy instrument. When a policy evaluation involves a thorough analysis of benefits and costs, this analysis may be limited at benefits and costs at the level of addressees (the first column), or it may consider all categories of costs presented in the table below.

	BODY AFFECTED	
TYPE OF COST	IMPLEMENTING BODY	ADDRESSEES
BUDGETARY COSTS	The direct financial outlays from the EU budget and other public funds. Moreover, they also include administrative costs for the Commission and public authorities, in particular the Member States, and human resources needed to manage the intervention.	Not applicable
TRANSACTION COSTS	Enforcement costs (associated with implementing, monitoring and enforcing the policy)	Costs incurred in identifying and selecting the most appropriate compliance route
COMPLIANCE COSTS:	Not applicable	These are the direct costs in complying with a policy measure. The resources used to covered compliance costs cannot be used for other purposes.
ADJUSTMENT COSTS	Not applicable	The costs of reallocating resources because of policy-induced changes in behaviour (production or consumption).

Allocative effects

As discussed above, spending programmes will normally provide benefits to the direct addressees and negative effects on other actors are usually not intended, even if they may appear. This contrasts with legislative measures which may in fact aim to prevent certain groups within society from continuing their usual and preferred behaviour in order to produce net benefits for society. Identifying the impacts of a public intervention on different groups in the society is hence crucial, if the intervention can be expected to have significant effects on such groups. In the policy formulation stage there are anticipated impacts on different social and economic groups, where the identification of "winners" and "losers" can be of help to foresee the obstacles to the proposal and may also point to a need to change the planned intervention. As the interventions implemented there are also likely to be impacts on existing inequalities. The availability of baseline data on, for example, geographical or social inequalities is essential to assess whether they are attenuated or aggravated by a policy over time.

Challenges of policy evaluation

As a reminder, below is a short list of some of the challenges of policy evaluation vis-à-vis programme evaluation which should always be kept in mind:

- The complexity of some policies vis-à-vis programmes in terms of objectives, instruments and addressees and evolutive nature,
- The need for administrations to demonstrate some sort of attainment, or at least progress towards the attainment of very high-level, longer-term objectives (i.e. often to shift the context in a significant way),
- Because of the latter, the complexity of causal relations given the multiplicity of levels of effects and interactions between the intervention and its effects, the context and other policies,
- The varied time scales for the concrete implementation of different types of instruments, and
- The differing nature of the expected cause and effect linkages between different instruments and their effects and the hence the differing rates of change that occur at the level of effects on addressees.

ANNEX G: QUALITY ASSESSMENT FRAMEWORK

The following is an example of the quality assessment framework in the Commission:

	UNACCEPTABLE	POOR	SATISFACTORY	GOOD	EXCELLENT
1.Meeting needs: Does the evaluation deal adequately with requests for information from the Commission and is it in line with the specifications?					
2. Relevant scope: Have the rationale of the programme, its outputs, results, impacts, interactions with other policies and unexpected effects been studied in full?					
3. Appropriate methodology: Is the design of the evaluation adequate and suitable for providing the results required (within time limits) to answer the main evaluation questions?					
4. Reliable data: Are the primary and secondary data collected or selected suitable? Are they sufficiently reliable in the light of the expected use?					
5. Sound analysis: Does the analysis of the quantitative and qualitative data comply with established rules, and is it complete and appropriate for answering the evaluation questions correctly?					
6. Credible results: Are the results logical and justified by the analysis of the data and by interpretations based on carefully presented explanatory hypotheses?					
7. Valuable conclusions: Are the conclusions just, and are they unbiased by personal or partisan considerations?					
8. Useful recommendations: Are the recommendations comprehensible, useful, applicable and detailed enough to be put into practical effect?					
9. Clarity: Does the report describe the context and goal of the programme evaluated and also the organisation and results in such a way that the information provided is easily understood?					

The use of this assessment grid is explained in section 4.2.3.

ANNEX H: POSSIBLE STRUCTURE FOR EX ANTE EVALUATION REPORT ⁸⁴

1. Introduction

- what is the purpose of this report?
- what is the structure of this report?
- what are the main sources of evidence and information on which it is based?
- what are the current Community expenditure or other interventions in this policy area?

2. What problem is the proposal expected to tackle?

- what is the problem in this policy area in terms of, for example social, economic and environmental terms?
- what are the underlying motive forces and who are the main actors involved?
- what is the concrete target group and what are its needs?
- who else would be affected?
- what would happen under a "no policy change" scenario?

3. What are the objectives that the proposal is expected to achieve?

- what is the overall policy objective in terms of expected impacts?
- has account been taken of other Community objectives?
- what are the general, specific and operational objectives and expected results?
- what indicators are planned for measuring inputs, outputs, results and impacts?

4. What are the main policy options and alternative delivery mechanisms?

- what is the basic approach chosen to reach the overall policy objective?
- what other policy instruments, besides an expenditure programme, were considered?
- what are the trade-offs associated with the proposed option?
- what other designs for an expenditure programme were considered?

5. Risks and assumptions

- what are the risks inherent in the current situation if nothing is done?
- what assumptions is the proposed expenditure programme based on?
- what risks are involved in implementing the proposed expenditure programme?
- what measures will be taken to counter these risks?

6. What positive and negative impacts are expected from the options considered?

- what are the expected impacts of each option (e.g. social, economic and environmental)?
- are there potential conflicts between different impacts?
- are there intense impacts on particular social groups, economic sectors or regions?
- are there impacts outside the EU on candidate and/or other countries?
- how are the impacts expected to manifest over time?
- what are the results of any scenario, risk or sensitivity analysis undertaken?

7. Added value of Community involvement

- how are subsidiarity and proportionality taken into account in the proposal?
- is the proposed programme complementary to other interventions?

⁸⁴ Adapted from the DG EAC model.

- how will synergies with other interventions be encouraged?
- how large are the additional ("marginal") effects that can be attributed to the proposal?

8. Stakeholder consultation and lessons learned

- which interested parties were consulted, and why? how was this done?
- what were the results of the consultation?
- what evaluations, audits, studies, or reports of similar interventions are available?
- what lessons and evidence have been taken into account in designing the proposal?

9. Helping to achieve cost-effectiveness

- what are the financial and human resource costs of the proposed programme?
- could the expected results be achieved at lower cost?
- could the same or better results be achieved at the same cost using other means?

10. Monitoring and evaluation

- what system will be in place to collect, store and process monitoring data?
- what types of evaluation will be carried out during the lifetime of the programme, and when?

11. Draft proposal and justification

- what is the final choice of option, and why?
- why was a more/less ambitious option not chosen instead?
- what are the trade-offs associated with the chosen option?
- what accompanying measures will be taken to maximise positive impacts and minimise negative impacts?

ANNEX I: GLOSSARY

Acceptability

The extent to which stakeholders accept the policy in general and the particular instrument proposed or employed.

Activity Based Budgeting (ABB)

Budgeting and financial management in an integrated manner on the basis of activities closely linked to the Commission's priorities and objectives. Provide the conceptual framework where resources are allocated and managed following pre-defined political priorities.

Activity Based Management (ABM)

Integrated management methodology covering all aspects of the Commission's work, including objective prioritisation, planning and programming, budgeting, management and reporting methods. It encompasses prioritisation and resource allocation at the level of the College and general principles and tools for management at the level of departments. It covers all resources managed by the Commission.

Action

Work carried out during a certain period of time, consuming resources and producing outputs and outcomes in accordance with the objectives set.

Activity

Coherent area of actions with objectives and resources, corresponding (or equivalent) to the ABB level. Each activity is related to only one policy area.

Allocative/distributional effects

The extent to which disproportionate negative/positive distributional effects of a policy are minimised/maximised.

Annual Activity Report (ARR)

Management tools that informs on the outputs and results of the management of a DG/Service for one year.

Annual Management Plan (AMP)

Results-oriented management and programming tool, translating the APS priorities into objectives while covering all the activities and resources of a DG/Service for a year.

APS

Annual Policy Strategy : Commission decision identifying political priorities, priority actions, and giving clear orientations on resources allocation.

Audit

A control function, which is primarily concerned with verifying the legality and regularity of the implementation of resources in a programme. Audit has traditionally covered areas such as the verification of financial records (financial audit). (see also 'performance audit').

Benchmarks

Standards by which the performance of an intervention can be assessed in a non-arbitrary fashion. An obvious way of deriving benchmarks would be to examine the intervention's objectives as expressed by expected outputs, results and outcomes. Ideally, benchmarks should allow us to compare the performance of an intervention with that of other policy instruments in the same field of action or in a related one.

Case studies

A data collection technique involving the examination of a limited number of specific cases or projects which the evaluator anticipates will be revealing about the programme as a whole. Case studies tend to be appropriate where it is extremely difficult to choose a sample large enough to be statistically representative to the population as a whole; where generalisation is not important; where in-depth, usually descriptive data is required; and where the cases or projects to be studied are likely to be quite complex.

Case study designs

A class of evaluation designs in the descriptive rather than the causal approach. It is often the case that an evaluation design will be based on an in-depth study of one or more specific cases or situations.

Coherence

The extent to which the intervention logic is non-contradictory/the intervention does not contradict other interventions with similar objectives.

Comparative change design

An example of a quasi-experimental design in which any known or recognisable difference between the programme and control groups is taken into account in the statistical analysis. The problems with this design are, firstly, that there may be some other factor which explains some or all of the variation in the intervention and in the observed effects, and, secondly, that there may be initial differences between the programme and control groups which have an influence on observed effects and which can therefore become confounded with the influence of the programme on these effects (selection bias).

Consistency

The extent to which positive/negative spillovers onto other economic, social or environmental policy areas are being maximised/minimised.

Control group

A group of subjects which have not been exposed to an intervention. The control group should resemble the programme group (the subjects which have been exposed to the intervention), so that systematic differences between the two groups may be attributed to the effects of the intervention once other plausible alternative hypotheses have been eliminated or discounted.

Cost-benefit analysis

A judgemental technique in which a researcher compares all social and private costs and benefits of an intervention with a view to determining whether the benefits exceed the costs, and if so by how much. Social costs and social benefits usually have to be measured by some indirect means and converted into monetary values so that a comparison can be

made with private costs and benefits. Furthermore, it may not be appropriate to use prevailing market prices. Consider a situation of very high unemployment. In this case, the real cost of labour may be much lower than the prevailing market wage. The opportunity cost (the next best use of the otherwise unemployed workers had the project not gone ahead) is lower than the prevailing wage rate, and this low opportunity cost has to be represented by a shadow price which has to be derived somehow.

Cost-effectiveness analysis

A judgmental technique in which the researcher quantifies the costs and benefits associated with an intervention on the basis of the same principles which apply to cost-benefit analysis, but there is no requirement to transfer benefits into common monetary units.

Counterfactual situation

The situation, which would have arisen had the intervention not taken place. In order to derive the counterfactual situation we need an evaluation design. Except for the theoretical case of the ideal experimental design, we can never know the counterfactual situation with certainty. Real world evaluation designs tend to be based on an estimate of the counterfactual derived either from comparing subjects who were exposed to an intervention with a comparison group who were not exposed, or from examining subjects before and after exposure.

Data

Known facts, which can be used as a basis for inference. Subjective data involve personal feelings, attitudes and perceptions; objective data relate to observable facts. Quantitative data involve numerical observations; qualitative data are non-numerical and related to categories. Longitudinal data are collected over time; cross-sectional data are collected from the same point in time, but from a variety of different geographical areas, etc. Primary data are taken directly from original sources or collected first hand; secondary data have undergone extensive manipulation and interpretation.

Data analysis

The main techniques used to interpret information about an intervention for use in an evaluation are statistical analysis, the use of models, non-statistical analysis and judgement techniques, such as cost-benefit analysis, cost-effectiveness analysis and multi-criteria analysis.

Data collection

The main techniques used to gather information about an intervention for use in an evaluation are surveys, case studies, natural observations, expert opinion, reviews of programme documents and literature reviews.

Deadweight

Deadweight is defined as effects, which would have arisen even if the intervention had not taken place. Deadweight usually arises as a result of inadequate delivery mechanisms, which fail to target the intervention's intended beneficiaries sufficiently well. As a result, other individuals and groups who are not included in the target population end up as recipients of benefits produced by the intervention. Deadweight is really a special case of programme inefficiency.

Delivery mechanisms

The organisational arrangements, which provide the goods and services funded by the intervention to its intended beneficiaries, i.e. its target population.

Delphi technique

A technique which can be used to systematise expert opinions. Experts are consulted separately in a number of different rounds. In each successive round, each individual is told the views of the other experts in the previous round. This technique can be used to arrive at a consensus, or at least to reduce disagreements.

Displacement

Displacement and substitution are two closely related terms which are used to describe situations where the effects of an intervention on a particular individual, group or area are only realised at the extent of other individuals, groups or areas. Consider, for example, the case of a programme to provide employment subsidies. In a firm, which benefits from this programme, subsidised workers may take the place of unsubsidised workers who would otherwise have been employed by that firm. This is known as substitution. Alternatively, a firm benefiting from the employment subsidies may win business from other firms, which do not participate in the scheme. Thus, the jobs created in the participating firm may be partly or wholly offset by job losses in other firms. This is known as displacement.

Dissemination

The set of activities by which knowledge about an evaluation is made available to the world at large.

Economy

The extent to which resources are available in due time, in appropriate quantity and quality at the best price. The Financial Regulation (article 27(2)) defines economy as the requirement that the resources used by the institution for the pursuit of its activities shall be made available in due time, in appropriate quantity and quality and at the best price.

Effectiveness

The extent to which objectives set are achieved. The Financial Regulation (article 27(2)) defines effectiveness as attaining the specific objectives set and achieving the intended results.

Efficiency

The extent to which the desired effects are achieved at a reasonable cost. The Financial Regulation (article 27(2)) defines efficiency as the best relationship between resources employed and results achieved.

Evaluability

The issue of whether or not the questions raised by a given analytical agenda for an evaluation are at all answerable by an evaluator using appropriate research methods. To know whether the questions can be answered with an acceptable degree of credibility, it is often advisable to perform an evaluability assessment. If an intervention is not evaluable in terms of this analytical agenda (e.g. because adequate data are not yet available), this can lead to a decision to postpone the evaluation or to draw up a new, more realistic analytical agenda.

Evaluation design

A model which is used to describe an intervention and provide evidence on the effects which may be attributable to it. Evaluation designs are either causal or descriptive in nature. A given design should lead to the choice of one or more data analysis and collection techniques.

Evaluation project

A sequence of logical steps starting out from the formulation of problems and interests motivating the evaluation to arrive at a series of questions that can be addressed in an analytically acceptable way. The aim is to establish a work plan setting out a framework in which the evaluation is to be conducted and then to choose the evaluator.

Evaluation report

The end product of an evaluation, the evaluation report must follow a logical structure and meet the needs of the evaluation sponsors and the principal stakeholders. Evaluation reports must include an executive summary of no more than five pages in length. The structure of the expected report is usually specified by the sponsors in the terms of reference.

Ex ante evaluation

An evaluation conducted before the implementation of an intervention. Also referred to as an "appraisal".

Ex post evaluation

An evaluation conducted either on or after completion of an intervention.

Expert opinion

A data collection technique, similar to a survey, which relies on the necessarily subjective views of experts in a particular field. It is not recommended to rely on expert opinion as a sole data source, for example, because of problems with so-called "chatty bias".

External evaluation

An evaluation which is performed by persons outside the organisation responsible for the intervention itself.

External validity

The confidence one can have about whether or not one's conclusions about the intervention can be generalised to fit circumstances, times, people, and so on, other than those of the intervention itself. A threat to external validity is an objection that the evaluation design does not allow causal inference about the intervention to be generalised to different times, places or subjects to those examined in the evaluation.

Feedback

The process by which the information compiled by an evaluation is used by decision-makers to either change the way in which an intervention is implemented, or to bring about a more fundamental change in the basic orientations of the intervention, including calling into question its very existence.

Formative evaluation

An evaluation concerned with examining ways of improving and enhancing the implementation and management of interventions. Formative evaluations tend to be conducted for the benefit of those managing the intervention with the intention of improving their work.

Formulation of judgements

Once all the data has been analysed and the evaluation results compiled the success of an intervention can be judged, a process which will focus in particular assessing an intervention in respect to specific evaluation questions that reflect generic evaluation issues such as effectiveness, efficiency, utility, sustainability and relevance.

Impacts

A general term used to describe the effects of an intervention on society. Impacts can be either positive or negative and foreseen or unforeseen. Initial impacts are called results, whilst longer-term impacts are called outcomes.

Indicator

A characteristic or attribute which can be measured to assess an intervention in terms of its outputs or impacts. Output indicators are normally straightforward. Impact indicators may be more difficult to derive, and it is often appropriate to rely on indirect indicators as proxies. Indicators can be either quantitative or qualitative. The term "performance indicators" is also used.

Inputs

The human and financial resources involved in the implementation of an intervention.

Interim evaluation

An evaluation conducted during the implementation of an intervention.

Internal evaluation

An evaluation, which is performed by members of the organisation responsible for the intervention itself.

Internal validity

The confidence one can have in one's conclusions about what the intervention actually did accomplish. A threat to internal validity is an objection that the evaluation design allows the causal link between the intervention and the observed effects to remain uncertain. It may be thought of as a question of the following nature: could not something else besides the intervention account for the difference between the situation after the intervention and the counterfactual?

Intervention

A generic term used to cover all public actions.

Intervention logic

The conceptual link from an intervention's inputs to the production of its outputs and, subsequently, to its impacts on society in terms of results and outcomes. The examination of the programme's intervention logic will be of central importance in most evaluations.

The evaluator needs to ask how the programme achieves its specific objectives, and how do the specific objectives contribute to the attainment of the general objectives? The terms "theory of action", "programme logic" and "programme theory" are sometimes used to mean more or less the same thing.

Monitoring

The continuous process of examining the delivery of programme outputs to intended beneficiaries, which is carried out during the execution of a programme with the intention of immediately correcting any deviation from operational objectives. Evaluation, on the other hand, is carried out at a discrete point in time, and consists of an in-depth study. Monitoring often generates data, which can be used in evaluations.

Multi-criteria analysis

A decision-making tool which can be adapted to form judgements about interventions. Multi-criteria analysis allows us to formulate judgements on the basis of multiple criteria, which may not have a common scaling and which may differ in relative importance.

Outcomes

The longer-term impact, usually expressed in terms of broad socio-economic consequences, which can be attributed to an intervention (e.g. a reduction in the number of long-term unemployed).

Outputs

The goods and services produced by an intervention (e.g. training courses for the long-term unemployed).

Performance audit

Conceptually closer to evaluation than traditional audit, performance audit is strongly concerned with questions of efficiency (of an intervention's direct outputs) and good management. Performance audit and evaluation share the same aim of improving the quality of programmes, but evaluation goes further. It also looks at issues such as sustainability, relevance and the longer-term consequences of a programme. (see also 'audit').

Policy

A set of activities, which may differ in type and have different direct beneficiaries, directed towards common general objectives. Policies are not delimited in terms of time schedule or budget.

Policy area

Coherent area of activities related to one Community policy or to an horizontal function and carried out by the Commission services.

Population

In statistics, the entire aggregate of individuals or subjects, from which samples can be drawn.

Quasi-experimental designs

A class of causal evaluation designs, which take a more practical approach than is the case with true experimental designs. Control groups can still be used, but these have to be assigned through some non-random process. Alternatively, one can examine beneficiaries before and after exposure to the intervention.

Regression analysis

A statistical inference technique which can be used to establish the significance of any correlation (association) between variables of interest, e.g. the gender of a long-term unemployed worker and the amount of time before he or she finds a new job after a training programme. In regression analysis, we attempt to establish whether the variation in one variable (known as the dependent variable) can be explained in terms of the variation in one or more independent variables. The dependent variable is often quantitative, e.g. a person's income can be regressed on his educational qualifications, number of hours worked per week, age, etc. Special techniques are available, however, to deal with situations in which the dependent variable is qualitative, e.g. whether or not a person owns a car can be regressed on income, wealth, age, gender etc.

Relevance

The extent to which an intervention's objectives are pertinent to needs, problems and issues to be addressed.

Reporting

Reporting takes place when the evaluator transmits the evaluation report (usually in the form of a document, or else through some audio-visual presentation) to the sponsors and when they, in turn, transmit a copy (or a summary thereof) to other interested parties.

Results

The initial impact of an intervention (e.g. an improvement in the employability of the long-term unemployed through a rise in their skill level).

Sample

A set of individuals or items selected from a given population so that properties and parameters of the population may be estimated, or so that hypotheses about that population may be estimated.

Scope

The field of investigation of an evaluation. Typically, this has to be defined from an institutional (EU versus national or local level), temporal (period review) and geographical (part of the EU territory) point of view. In addition, one has to identify the key evaluation issues (relevance, efficiency, effectiveness, utility, sustainability) which will be examined.

Stakeholders

The various individuals and organisations who are directly and indirectly affected by the implementation and results of a given intervention, and who are likely to have an interest in its evaluation (e.g. programme managers, policy-makers, the programme's target population).

Statistical analysis

A commonly used data analysis technique. Statistical analysis is often used to describe phenomena in a concise and revealing manner. This is known as descriptive statistics. It can also be used to test for relationships among variables or generalise findings to a wider population. This is known as statistical inference.

Steering group

Part of the management structure for an evaluation, a steering group allows other services (and possibly other stakeholders from outside the Commission) to contribute to the development of the evaluation project. See also evaluation project, management structure, stakeholders.

Strategic Programming and Planning

Continuous and systematic process whereby the Commission translates in a cycle comprising several steps involving its services in strategic decisions on objectives, activities and resources.

Structuring phase

Covers the period from the drawing up of the terms of reference by officials to the delivery of the inception report by the evaluator. In this phase the effects to be evaluated are chosen and classified and the observation tools are clearly defined. The structuring phase embraces preparatory work, drafting the terms of reference and it normally culminates with the inception report.

Summative evaluation

An evaluation concerned with determining the essential effectiveness of an intervention. Summative evaluations tend to be conducted for the benefit of external actors (groups who are not directly involved in the management of a programme), for reasons of accountability or to assist in the allocation of budgetary resources.

Surveys

A widely-used technique for collecting data from a sample drawn from a given population. Surveys are often based on probability sampling, and survey information is usually obtained through structured interviews or self-administered questionnaires. Cross-sectional surveys involve measurements made at a single point in time. Panel surveys involve measurements acquired at two or more points in time.

Sustainability

The extent to which positive effects are likely to last after an intervention has terminated.

Target population

The intended beneficiaries (individuals, households, groups, firms) of an intervention. An intervention may have more than one target population. This term should be distinguished from "population" in the statistical sense.

Terms of reference

The terms of reference outline the work to be carried out by the evaluator, the questions to be dealt with and the time schedule. They allow the sponsors of the evaluation to define their requirements and allow the evaluator to understand clearly what is expected of the

work to be undertaken (including, often, the structure of the expected evaluation report). Clearly defined terms of reference are vitally important where an evaluation is to be conducted by an external expert, and can also be of tremendous use when it is to be performed in-house.

Thematic evaluation

An evaluation, which focuses on one or more themes, which are common to several different interventions (programmes or other activities), for example, effects on the environment or on small and medium-sized enterprises.

Utility

The extent to which effects correspond with the needs, problems and issues to be addressed.

Work Programme

Management tool implementing annual work programme defining the Commission's major political priorities and identifying concrete actions, legislative or executive initiatives that translate those priorities into operational terms.

European Commission

Evaluating EU Activities — A practical guide for the Commission services

Luxembourg: Office for Official Publications of the European Communities

2004 — 110 pp. — 21 x 29,7 cm

ISBN 92-894-7928-0

