



## **AUDIT GUIDELINES ON EVALUATION**

### **TABLE OF CONTENTS**

#### **FOREWORD**

#### **SECTION 1: Introduction to evaluation**

- 1.1. What is Evaluation ?
- 1.2. Evaluation at the Commission

#### **SECTION 2: Relevance of evaluation for Performance audit**

- 2.1. Using elements of evaluations to inform Performance audits
- 2.2. Auditing individual evaluation
- 2.3. Assessing the quality of an evaluation
- 2.4. Auditing the evaluation system of a Directorate [General \(DG\)](#)

#### **SECTION 3: Sources of further information**

- 3.1. Evaluation in general
- 3.2. Evaluation at the Commission
- 3.3. Court's audit relating to evaluation

### **FOREWORD**

The development of performance audit at the European Court of Auditors and the increasing use of evaluation by the Commission means that evaluations of EU activities are becoming increasingly relevant to the work of the Court.

The Court's Audit Manual outlines the relationship between the two as follows: "Performance audits will usually include evaluative elements of selected subjects and consider evaluation systems and information with a view to assessing their quality and, when they are considered to be satisfactory and relevant, use evaluation information as audit evidence".

The aim of the document is to provide auditors with an overview of:

- what evaluation is and how it is used and managed by the Commission (section 1);
- how evaluation can be assessed and used as audit evidence in Performance audit (section 2).

More detailed guidance and suggestions are provided in Appendix A (Key elements for building up a programme for auditing an evaluation) and Appendix B (Key elements for building up a programme for auditing an evaluation system). Sources of further information are included at the end (section 3).

#### **List of documents related**

**Appendix A: Key elements for building up a programme for auditing an evaluation**

**Appendix B: Key elements for building up a programme for auditing an evaluation system**

#### **Who to contact**

If you feel that the information provided in this document could be improved, please do not hesitate to communicate your suggestions: [ECA-AMS.CONTACT@eca.europa.eu](mailto:ECA-AMS.CONTACT@eca.europa.eu)

## SECTION 1: INTRODUCTION TO EVALUATION

### 1.1. WHAT IS EVALUATION?

#### Evidence-based judgement of the value of public interventions

Many different types of appraisal or assessment activities are referred to as *evaluations*. This guideline applies to formal evaluation carried out to reach judgements about the actual or likely effects of public interventions such as expenditure programmes, financial instruments and regulatory measures.

Formal evaluation has been developing since the 1950's and is now carried out routinely in many public administrations across the world, including the European Commission. Although there is no single agreed definition of evaluation, there is a general consensus that evaluation is a process that makes evidence based judgements about public interventions. These judgements are formed on the basis of a disciplined inquiry involving procedures which draw heavily on techniques from the fields of social science and public policy for collecting and analysing quantitative and qualitative information.

#### A range of criteria and standpoints can be applied

Evaluations usually address questions based on one or more of the following types of criteria:

- the **relevance** of interventions to the needs they aim to satisfy e.g. *Did the intervention address the priorities of the target group?*;
- the **economy, efficiency, effectiveness** with which results and impacts are produced e.g. *What were the costs-benefits to budgets and/or stakeholders?; What measures could be taken to improve delivery? Did the intervention achieve its objectives?; and*
- the **sustainability** of the effects on interventions e.g. *What results will persist after the intervention has ceased?*

In doing this, evaluations frequently examine the cause and effect, the assumptions underlying interventions and the political objectives.

Reaching judgements about intervention in terms of the criteria above involves adopting the standpoint of one or more of the stakeholders in the intervention; for example, a question about the "economy" of an intervention may be framed in terms of its budgetary cost (i.e. adopting the standpoint of the budgetary authority), or in terms of the costs to businesses and citizens (i.e. adopting the standpoint of those upon whom the impact of the intervention falls). For this reason evaluation processes usually include provision for stakeholder involvement. This is frequently achieved through the use of "evaluation committees" or "steering groups" made up of stakeholders or their representatives who meet regularly with evaluator during the evaluation process to help facilitate the work of the evaluator.

#### The information can have many users and uses

The information produced during an evaluation is discussed amongst the participants in the evaluation process and is often captured in a series of reports culminating in a final report containing the findings, results, conclusions and recommendations. The report, or its content in a suitably amended format, is then communicated or made available to intended or potential users.

Evaluation is thus a key source of information for those within an organisation responsible for managing the implementation of interventions or the developing new policies. These stakeholders are frequently involved in the evaluation process itself and are able to feedback some evaluation results directly without waiting for the finalisation of the report.

Equally evaluations can provide information to those responsible for setting political priorities. However, it is always stressed in the evaluation community that evaluation is no substitute for political judgement; it is one source of information available to political decision makers that has to be balanced against others.

In addition, public sector organisations also aim to use evaluations cumulatively to improve decision making and acquire knowledge and organisational learning (e.g. by creating a stock of evaluation results, setting up internal networks of users and/or practitioners and performing secondary analysis or meta-evaluation on evaluations carried out).

**A great variety of evaluations are carried out**

As a result of the relatively recent development of evaluation as a discipline and the open nature of the community of evaluation practitioners, evaluation studies can vary greatly from one another depending on the context in which they are carried out and choices made within the evaluation process.

The main elements that make up the context of an evaluation are:

- subject**
  - **what the subject of the evaluation is** - it could be a project, a programme, a regulatory measure, a financial instrument, a policy, a theme or even an organisation.
- evaluation & management system**
  - **what the relationship between the evaluation and the management system is** - in some management systems regular evaluation is an extension of performance monitoring and reporting arrangements in other cases it is used mostly on an ad hoc basis to address specific questions or issues that have arisen;
- sponsor**
  - **who is sponsoring or carrying out the evaluation** - the organisation sponsoring or carrying out the evaluation could either be evaluating its own activities and achievements or those of other organisations. In addition, the evaluation could be carried out by specialists from within the organisation, mainly by external professionals, or a mixture of the two. Unlike performance audits, where objectivity is ensured by the organisational independence of auditors, objectivity and lack of bias tends to be assured through methods and stakeholder participation rather than the organisational independence of the evaluator.
- funding**
  - **how it is funded** - e.g. funding could come from within a programme being evaluated and thus be controlled by programme managers or from funds set aside for evaluation with commissioning and control of funds independent of the management of the programme.
- purpose**
  - **what the purpose is** - a number of purposes are possible and purposes could be combined within a single evaluation. The most commonly cited purposes are to aid decision making about interventions, the setting of political priorities, and resource allocation and to contribute to accountability and organisational learning.
- users**
  - **who the intended users are** - evaluations can address issues or questions of relevance to one or more stakeholders of the intervention (e.g. managers, political authorities, the addressees of the intervention or the wider public).
- prospective or retrospective**
  - **whether it is prospective or retrospective** - evaluation is mostly carried out retrospectively (i.e. after or during the implementation of an intervention) to assess its outcomes, but is increasingly being carried out prospectively to provide assessments of the likely outcomes of proposed interventions (e.g. regulatory measures or expenditure programmes).

The elements of the context of the evaluation listed above are usually outside the control of the evaluation process itself. Nevertheless, they should be considered by evaluation project managers when making decisions about the elements of the evaluation process that largely determine its quality i.e.:

- the **management arrangements**;
- the identity of the **evaluator**;
- the **participants** in the process;
- the **issues** addressed i.e. the questions asked;
- the **methods employed** for collecting and analysing data, and
- the **timetable and arrangements for reporting the results**.

## 1.2. EVALUATION AT THE COMMISSION

### **A wide range of activities is evaluated**

EU activity is itself wide ranging covering many different policy areas. The types of intervention used for achieving EU objectives also vary considerably between the policy areas.

The result is that evaluations cover the whole gamut of potential subjects i.e. individual projects within programmes, whole programmes, regulatory measures, financial instruments, whole policies involving a mixture of programmes and other measures, cross-cutting themes and even Community agencies.

### **Monitoring arrangements vary considerably between interventions**

The role of the Commission in a particular area and the management method applicable largely determine the monitoring arrangements in place and the role of evaluation. The monitoring arrangements and the nature and type of monitoring data available have a significant influence on the type of evaluation carried out e.g. in some evaluations data collection is an important part whereas in other cases evaluations are based on pre-existing data with the emphasis more on analysis or meta-evaluation<sup>1</sup>.

### **The Commission is the major but not exclusive sponsor of evaluations of EU activity**

The Commission is the sponsor of most evaluations of EU activity. However, notable exceptions exist such as the ex ante and interim evaluations in the Structural Funds (which are the responsibility of the individual Member States) and recipients of grants or project funding (who are required to have their activities evaluated as part of the conditions for receiving financing).

In some areas the Commission carries out evaluations of activities which it is directly responsible for managing (e.g. its own administrative activities) but in most domains the Commission evaluates the effects of activities that are implemented by others but for which the Commission has an overall responsibility.

### **Funding arrangements depend on the type of intervention**

Almost all evaluations of EU expenditure are funded through the budget. In the case of expenditure programmes, provision is usually made within the regulations governing them for funding evaluation out of programme appropriations. In other cases evaluations are funded out of Directorate Generals (DG) administrative budgets. It can also be the case that evaluations of the effects of EU policies are carried out independently of the Commission by Member States themselves and accordingly funded by them.

### **Evaluations usually serve more than one purpose**

The evaluations of the Commission can cover a wide range of purposes and be addressed to a wide range of users. The Commission describes the four main types of purpose for evaluations as:

- 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).

These types of purpose are frequently combined within the same evaluation exercise. This is particularly likely to be the case for evaluations carried out of whole programmes, policies or cross-cutting themes.

### **Evaluations usually address multiple users and are used cumulatively**

Often the needs of different users are addressed simultaneously. The main users are the officials of the Commission responsible for managing or proposing interventions. To a lesser extent evaluations are carried out specifically to help the College of Commissioners with the setting of political priorities. These are known at the Commission as "strategic" or "cross-cutting" evaluations and address issues relevant to more than one activity or policy area e.g. support for small and medium sized enterprises (SMEs). They feed directly into the process for establishing the Annual Policy Strategy of the Commission which also involves discussion with the Council and the Parliament, who are also significant users of evaluation results in the budgetary, discharge and legislative procedures.

<sup>1</sup> Synthesis based on a series of evaluations. The term is also used to refer to an evaluation of another evaluation or of a series of evaluations.

## Both prospective and retrospective evaluations are carried out

The Commission sponsors both prospective and retrospective evaluations. The main types of retrospective evaluation are known as mid-term or interim and ex post evaluations.

Two types of prospective analysis are carried out: ex ante evaluations and impact assessments. Ex ante evaluations are carried out on proposals that are likely to have a significant budgetary impact and include judgements of the potential cost-effectiveness of the proposed intervention. Impact assessments involve assessments of the potential social, economic and environmental impacts of items in the Commission's work programme. There is considerable overlap between them and in practice impact assessments tend to be combined in a single exercise with ex ante evaluation where both are required.

Similarly, in practice, evaluations often combine ex ante, interim and ex post elements e.g. the evaluation of an ongoing expenditure programme may combine ex post evaluation of elements of the previous programme with interim evaluation of the current programme and ex ante evaluation of elements of the proposal for the next programme.

## There are many evaluation systems at the Commission which operate within a wide framework

The Commission has over time developed evaluation systems specific to each area. As a result, the overall approach to evaluation can vary significantly according to the policy area and between DGs. Partly in response to this diversity and in order to encourage more widespread use of evaluation, the Commission has been since mid-1990s developing an evaluation policy that applies to all DGs and provides a framework within which DG's individual evaluation systems must operate.

Evaluation is seen as an integral part of the monitoring arrangements of the Commission's Activity Based Management initiative and acknowledged in the Financial Regulation as part of the Commission's internal control system for ensuring the sound financial management of EU funds. In addition, evaluation, particularly in the form of impact assessment, has been allocated a key role in the Commission's efforts towards "Better Regulation"<sup>2</sup>.

The Commission's evaluation policy is based around the requirement for DGs to evaluate their activities regularly and a set of standards governing their evaluation arrangements.

## Requirements to evaluate

In brief, DGs are required to evaluate their activities which have addressees external to the Commission. New activities should be the subject of a prospective evaluation which usually takes the form of an impact assessment, as is required for all items in the Commission's work programme, or an ex ante evaluation. Activities should be evaluated retrospectively at least once every six years with multi-annual expenditure programmes requiring at least one thorough evaluation during their life cycle. There is also a requirement to evaluate pilot projects and preliminary actions where these are likely to be followed by proposals for expenditure programmes. "Strategic" or "cross-cutting" evaluations are also carried out as required by the College related to the political priorities and strategic objectives of the Commission which involve more than one activity or policy area.

## Common features of the different evaluation systems

Most DGs<sup>3</sup> are required to have evaluation systems with a number of features in common as required by the Commission's Evaluation Standards and Good Practices. The main features are:

- **Evaluation planning** - DGs are required, as part of their Annual Management Plans, to draw up an annual evaluation plan with a multi-annual component which ensures that requirements to evaluate are met and that enough information from evaluations is available in time to support decision making.
- **Evaluation functions** - most DGs have a unit or sector specialising in carrying out or supporting evaluations. This unit is also responsible for establishing the evaluation plan, disseminating / reporting on the results of evaluations and developing methods, guidance and training.
- **Project management of evaluations** - one official, from either the evaluation function or one of the units whose activities are being evaluated, is normally responsible for the management of the evaluation. This involves drafting the mandate/terms of reference, establishing a steering group, administering the procurement procedures, liaising with the external evaluator, and leading the assessment of quality, and disseminating the results.
- **Steering groups** - a group of people chosen to facilitate the evaluator's access to data, provide methodological support, and take part in quality assessment. Stakeholders are usually included e.g. representatives of other DGs and services, Member State officials or representatives of the addressees of the intervention. Steering groups are not obligatory but in practice about 80% of programme/policy level evaluations have them.
- **Quality assessment** - DGs should have established quality standards that are incorporated in the terms of reference and which form the basis of the regular assessments of the evaluation work by the project manager and steering group. A formal quality assessment in writing of the draft and final reports is usually produced.
- **Procedures for following up results** - DGs should have procedures for ensuring that: operational units respond to evaluation findings, results, conclusions and recommendations; action plans are drawn up where appropriate; and, their implementation is monitored by senior management.

<sup>2</sup> [http://ec.europa.eu/governance/better\\_regulation/index\\_en.htm](http://ec.europa.eu/governance/better_regulation/index_en.htm)

<sup>3</sup>DGs whose activities have addressees external to the Commission have to evaluate their activities and therefore should have an Evaluation Function.

## SECTION 2: RELEVANCE OF EVALUATION FOR PERFORMANCE AUDIT

### When evaluation is relevant for audit

There are three main circumstances in which evaluations or evaluation-related activities can be relevant to an audit:

1. when **the evaluation provides information relevant** to one or more of the phases of the audit, e.g. programme descriptions, baseline data, the conclusion of a particular piece of analysis or even a recommendation (section 2.1.);
2. when **auditing the individual evaluation is relevant** to one or more of the objectives of the audit, e.g. the evaluation of the effectiveness of a particular project or measure (section 2.2.);
3. when **the evaluation system is the subject of the audit** i.e. a collection of evaluations and the processes for selecting and managing them and using the results are being assessed (section 2.3.).

### 2.1. USING ELEMENTS OF EVALUATIONS TO INFORM PERFORMANCE AUDITS

#### Ideas at the programming and planning stage

Many elements of individual evaluations or related activities are potentially good sources of information capable of providing ideas at the programming and planning stages or evidence and recommendations during fieldwork and reporting.

The Commission's evaluation activities and the wider evaluation community are potentially a good source of ideas relevant to selecting and planning audits:

- DGs' **evaluation plans and programmes** give insight into upcoming decisions on policies, programmes and political priorities as well as details on the timing and purpose of planned evaluations. This information could be useful to auditors in identifying audit subjects or objectives;
- **manuals and guidance** can provide useful explanations on the use of the many techniques common to evaluation and performance audit - individual DGs, Budget DG, SAs, international organisations and evaluation societies all produce materials that are potentially useful to auditors in developing their approach and methods;
- the Commission is a major user of external expertise maintaining **lists of experts** with quality assessments of the work of different providers published by Budget DG on the intranet. Most providers are also associated to evaluation societies. Such information is helpful to auditors seeking to use the services of external experts both in identifying suitable candidates and possible conflicts of interest;
- **Evaluation reports** (including impact assessments) usually include a considerable amount of information that will enable auditors to understand the audit field and plan the audit e.g. to identify relevant issues or problems, the stakeholders, the objectives of interventions, the logic of the intervention, potential audit criteria, the data available and appropriate methods for analysis.

#### Potential source of audit evidence and recommendations

There are a number of ways in which information created by evaluation activities could be used during the fieldwork and reporting stages of the audit:

- Information collected from past evaluations can act as **baselines** for performance audits e.g. in audits aiming to conclude on whether initiatives taken resulted in improvements;
- Evaluations often involve the **creation of information** not available from other sources that could either be used directly or reanalysed to address audit questions;
- Good quality evaluations can produce **reliable conclusions** about the economy, efficiency or effectiveness of an intervention directly relevant to one or more of the audit objectives;
- Evaluations results can also act as **corroborative evidence** for the Court's own studies, e.g. where the audit was carried out in parallel on a different basis but address some of the same issues, in such a case the results of the evaluation could either provide further positive evidence supporting the Court's conclusions or where different conclusions were reached lead auditors to reconsider their evidence and if necessary carry out further testing;
- Evaluation reports and responses to them usually contain **recommendations** or action plans which may or may not have been acted upon - reviewing recommendations and their follow up can thus provide evidence about the quality of management as well as give examples of pertinent recommendations not implemented which could be taken up by the Court.

Where the findings, results or conclusions of an evaluation are relevant to the objectives of audit they may only be used as audit evidence after carrying out appropriate testing to assess their reliability or validity. Such testing will need to focus on the **robustness** of methods for data collection and analysis used, the **coherence** between the data set collected the analysis used and the conclusions drawn, and the **impartiality** or objectivity of participants with respect to the judgemental aspects of the evaluation process (see section below).

## 2.2. AUDITING INDIVIDUAL EVALUATION

### When to audit an individual evaluation

There are three main circumstances in which the Court may wish to audit an individual evaluation project:

1. the evaluation is highly relevant to the Parliament and Council and likely to be a significant input to decisions about substantial amounts of EU funds,
2. as part of audit testing designed to gather evidence about quality of the evaluation system of a DG or the management of the activity evaluated;
3. as part of testing to gather audit evidence directly about the effects of projects, programmes or other measures/instruments where the objectives of the evaluation overlap considerably with those of the audit.

## 2.3. ASSESSING THE QUALITY OF AN EVALUATION

### Quality elements to assess

The main qualities that an evaluation should possess are:

- **usefulness** - the evaluation should address issues that are relevant to intended users / decisions that it is intended to support, be delivered on time and in an appropriate manner;
- **coherence** - the methods for collecting and analysing data should be appropriate for answering the questions, the human and financial resources for conducting the evaluation should be sufficient, and the management structure should be appropriate;
- **robustness** - the data collection and analysis should be carried out rigorously so as to accurately reflect the underlying reality;
- **impartiality** - the reported conclusions should be free from bias and fairly reflect evaluation findings;
- **clarity** - the documents making up the evaluation should be clear about the context and purpose of the evaluation, the questions addressed, assumptions made (e.g. the limits of data collected or analytical techniques used), data collected, methods used, the results obtained, the conclusions drawn and recommendations offered;
- **cost effectiveness** - the evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified.

Obviously trade-offs can exist between these qualities, e.g. pressure to deliver results in time for decisions (usefulness) may lead to a less accurate but less time consuming methods being employed (robustness). The qualities can also be mutually reinforcing either positively or negatively, e.g. a lack of coherence in the design of the evaluation is likely to lead to a lack of clarity in the final report.

Judging the quality of an evaluation is about assessing whether trade-offs made were appropriate in the circumstances and did not unduly compromise other aspects of the evaluation. Having less robust but more timely results can be justified if it contributes to the usefulness of the evaluation so long as methodological compromises are properly taken into account in arriving at conclusions (i.e. coherence is not compromised) and the limits of the analysis are clearly stated in the report (i.e. clarity maintained). Thus the evaluation is likely to be cost effective; the alternative, an evaluation producing reliable results too late would be worthless.

### Audit testing

Depending on the specific audit objectives testing could focus on one or more of these qualities. Assessing the extent to which an evaluation possesses the qualities above should be done on the basis of testing related to:

- the management of the evaluation process,
- the evaluation report, and
- underlying data and analysis.

### More practical guidelines

Practical guidelines and suggestions for building up a programme for auditing an evaluation are provided in Appendix A.

## 2.4. AUDITING THE EVALUATION SYSTEM OF A DG

Evaluation needs can vary widely between policy areas depending on the nature of the activities carried out. Evaluation systems should therefore be adapted to circumstances taking due account of the costs involved.

### Commission standards

A good evaluation system should have measures in place to manage issues related to the demand, supply and use of the evaluations that are crucial to ensuring that sufficient good quality evaluations are produced and used to support decision making and foster organisational learning. To some extent the Commission's Evaluation Standards provide a framework or "blueprint" that applies to all DGs. However, the standards leave considerable room for DGs to make choices about how they should be implemented. Listed below are a number of key qualities that an evaluation system should strive to achieve in order to ensure that it produces and uses sufficient quantities of good quality evaluations.

### Getting demand right

Effective management of the demand for evaluation is about:

- guaranteeing **support from the "top"** for evaluations i.e. from high level decision makers, those responsible for allocating resources or from managers of interventions e.g. through agreement of the evaluation mandate by the Director General or involving operational directorates in the process for selecting evaluations and showing commitment to use evaluations;
- creating **reasonable expectations** on the part of stakeholders about what evaluation can deliver - unreasonable expectations will only be disappointed leading to disillusionment which undermines support for and use of evaluation e.g. through evaluation training for senior staff or involving stakeholders or their representatives in steering groups;
- having the **right mix of incentives** for carrying out evaluations in terms of "sticks, carrots and sermons" e.g. requirements to evaluate (stick), earmarked funds (carrot) or encouragement through awareness raising exercises such as training (sermons);
- providing **sufficient links with decision making** processes e.g. regarding renewal decisions, setting priorities or allocating resources that help create an expectation for evaluative material and an outlet for its use.

### Ensuring the quality of the supply

Having the capacity to produce sufficient quantities of good quality evaluation efficiently requires DGs to have policies, arrangements or procedures for:

- recognising the needs of staff for **training and support** in their respective roles in the evaluation process and meet them appropriately, e.g. training and support targeted specifically at project managers, steering group participants, and programme managers;
- ensuring **complementarity with monitoring and audit systems** where this would be cost effective, e.g. procedures for ensuring that data collection, monitoring and evaluation arrangements are built into proposal for new interventions and that there is no overlap with auditing;
- **planning evaluations** so as to ensure that, wherever feasible or useful, evaluation results are available to support decision making and to maximise the chance of them being of good quality, delivered on time and in budget, e.g. decisions need to be anticipated, resources mobilised and evaluations launched promptly;
- **involving stakeholders** in an appropriate way so as to provide access to data, support the work of the evaluator from a methodological viewpoint, contribute to the assessment of the evaluation's quality and to encourage use of the results without unduly compromising the cost, management, and timing of the report or the credibility of the results, e.g. through the use of steering groups.
- **ensuring methodological quality** (i.e. robustness and coherence) in the execution of the evaluation without unduly compromising its usefulness or cost, e.g. through applying quality control procedures at the inception stage.
- **choosing appropriate evaluators** with sufficient knowledge or experience of both evaluation and the policy area. This goes beyond simply administering public procurement procedures, as steps often have to be taken over time in order to grow the supply of external expertise capable of carrying out evaluations of EU policies, e.g. by publicising the need and the resources available for commissioning studies



## Encouraging the use of results

For evaluations to be used it is not sufficient just to create demand for them and ensure that evaluations carried out are of good quality, in addition policies, arrangements and/or procedures should be in place for:

- **identifying users and their needs** as an input to the processes for determining the focus of the evaluation and the strategy for disseminating the results;
- **ensuring questions are relevant** to users and in line with the overall purpose of the evaluation;
- **making judgements and recommendations** that are credible and thought provoking as part of the evaluation process thus encouraging discussion and follow up.
- **communicating findings** in ways that are appropriate to intended users to maximise their impact;
- **delivering evaluations in time** for them to be used or where this is not possible for ensuring that early findings are disseminated or even that a halt is called to the evaluation;
- **monitoring use and follow up** of findings, results, conclusions and recommendations to provide feedback on what kind has the most impact.

## More practical guidelines

Practical guidelines and suggestions for building up a programme for auditing an evaluation system are provided in Appendix B.

# SECTION 3: SOURCES OF FURTHER INFORMATION

## EVALUATION IN GENERAL

Evaluation is a domain that is constantly evolving with developments being made by public administrations, international organisations and private firms all over the world. These bodies come together in evaluation societies to share information. The websites of these societies provide a good deal of evaluation related material in a variety of languages:

### European National Organisations

- European Evaluation Society (<http://www.europeanevaluation.org>) (including links to international evaluation sites)
- Italian Evaluation Association (<http://www.valutazioneitaliana.it/>)
- German Evaluation Society (<http://www.degeval.de/>)
- French Evaluation Society (<http://www.sfe.asso.fr/>)
- UK Evaluation Society (<http://www.evaluation.org.uk/>)
- Danish Evaluation Society (<http://www.danskevalueringsselskab.dk/>)
- Swiss Evaluation Society (<http://www.seval.ch/en/index.cfm>)
- Polish Evaluation Society (<http://www.pte.org.pl/>)
- Swedish Evaluation Society (<http://www.statskontoret.se/utvarderarna/hem.html>)
- Finnish Evaluation Society (<http://www.finnishevaluationsociety.net/>)
- Russian Evaluation Network (<http://ipen21.org/ipen/en/default.html>)
- Wallon Evaluation Society (<http://www.prospeval.org/>)
- Spanish Public Policy Evaluation Society (<http://www.idr.es/>)

### International Organisations

A number of international organisations also maintain evaluation websites:

- OECD Development Assistance Committee (<http://www.oecd.org/dac>) - including development cooperation reviews (including with the European Commission), information on high level performance indicators, details of publications and more links
- UN Development programme (<http://www.undp.org>) is a portal to the Evaluation Office function within the UNDP and includes evaluation reports, publications, documentation about the evaluation function in UNDP, its evaluation agenda, a search database and links to other evaluation websites
- UNICEF ([http://www.unicef.org/evaluation/index\\_evaluation.html](http://www.unicef.org/evaluation/index_evaluation.html)) gives an introduction to evaluation, lists the evaluation policy in the UNICEF, has an evaluation database, identifies good practises as a part of the evaluation process and provides links to other evaluation websites
- World Bank Operations Evaluation Department (<http://www.worldbank.org/evaluation>) Information about the World Bank's evaluation systems, and copies of evaluation reports

### Supreme Audit Institutions

- General Accounting Office (<http://www.gao.gov/policy/guidance.htm>) (USA): guides on evaluation and audit



## AUDIT GUIDELINES ON EVALUATION - APPENDIX A

# KEY ELEMENTS FOR BUILDING UP A PROGRAMME FOR AUDITING AN EVALUATION

### TABLE OF CONTENTS

#### INTRODUCTION

#### SECTION 1: Context, management arrangements and risks

Context of the evaluation

The management of the evaluation process

Risk to the quality of the evaluation

#### SECTION 2: Audit testing

Gathering preliminary information

Audit techniques

### INTRODUCTION

This note provides practical guidelines on how to assess the quality of an individual evaluation. As evaluations vary considerably in terms of the context in which and how they are managed, it is not practicable to develop a "one size fits all" audit programme. However, the following steps should help auditors at the European Court of Auditors to exercise their professional judgement in a consistent manner when developing programmes for auditing evaluations:

- a) analyse the context of the evaluation and the management arrangements related to it;
- b) identify the main risks to quality;
- c) develop an appropriate audit programme containing audit tests to address the areas of risk identified;
- d) carry out the planned testing and analyse the results.

**Section 1** relates to steps a) and b). It explains the main elements that make up the context and management arrangements of an evaluation. Then, under each component of quality, it gives the main types of risk normally associated with them and provides a list of common risk indicators.

**Section 2** relates to steps c) and d). It outlines the main audit techniques that are mostly likely to be appropriate to address the risks identified.

#### List of documents related

#### Audit Guidelines on Evaluation

**Appendix B: Key elements for building up a programme for auditing an evaluation system**

#### Who to contact

If you feel that the information provided in this document could be improved, please do not hesitate to communicate your suggestions: [ECA-AMS.CONTACT@eca.europa.eu](mailto:ECA-AMS.CONTACT@eca.europa.eu).

## SECTION 1: CONTEXT, MANAGEMENT ARRANGEMENTS AND RISKS

Evaluation studies can vary greatly from one another, depending on the context in which they are carried out and choices made within the evaluation process. A key aspect of judging the quality of an evaluation will thus be to assess the extent to which the evaluation was appropriate to its context. To do this one must first establish the context and how the evaluation was managed.

### CONTEXT OF THE EVALUATION

Identifying the context of the evaluation is about determining:

<b>subject</b>	<b>the subject</b> - What is being evaluated? Is it a project, a programme, a regulatory measure, a financial instrument, a policy, a theme or an organisation? What time period is being covered?
<b>methods and arrangements</b>	<b>implementation methods and monitoring arrangements</b> - What is the Commission's responsibility in the policy area? Is the budget implemented on a centralised basis, by shared or decentralised management or by joint management with international organisations? What kind of monitoring systems are in place? What sort of monitoring information was available? What use of existing management information was foreseen by the evaluation?
<b>sponsor of the evaluation</b>	<b>the sponsor of the evaluation</b> - Was the evaluation commissioned or carried out by the organisation responsible for the activity evaluated? Was it carried out at the request of another organisation or required by law / rules?
<b>funding arrangements</b>	<b>funding arrangements</b> - Was the evaluation funded from the same funds as the activity and controlled by the managers of the activity evaluated or by someone else?
<b>purpose</b>	<b>the purpose</b> - What purposes did the evaluation serve, e.g. aiding decision making about interventions, the setting of political priorities, resource allocation, accountability and/or organisational learning?
<b>intended users</b>	<b>the intended users</b> - Was the evaluation designed to address issues or questions of relevance to one or more stakeholders? Who were they e.g. managers, political authorities, the addressees of the intervention or the wider public?
<b>prospective or retrospective?</b>	<b>whether it is prospective or retrospective</b> - Was the evaluation designed to assess the actual or potential outcome of the activity / intervention? Was the evaluation carried out before, during or after implementation?

### THE MANAGEMENT OF THE EVALUATION PROCESS

Documenting the evaluation process is about identifying:

<b>management arrangements</b>	the <b>management arrangements</b> - Was the evaluation set up as a project? (Who was the project manager? What was his position in the organisation, experience, and other responsibilities?), Was a steering committee set up? (Who did it consist of? What was their experience? What role did it play? What meetings were held?), How was the process documented? (Was there a mandate, terms of reference, inception / interim / draft final and final reports?) Was there a procurement procedure? (How long did it take? What value was the contract? What criteria were used to assess and select offers? What tenders were received? Who evaluated the tenders?);
<b>evaluator</b>	the <b>evaluator</b> - Was the evaluator internal or external? What experience did they have of evaluation / the policy area? What qualifications did they have? Had they previously worked for the Commission? What other clients do they have?
<b>participants</b>	the <b>participants</b> in the process - Who was involved in the evaluation (in addition to the project manager, steering committee and tender committee)? What role did senior management play? Was the evaluation function involved? Were intended users engaged in the process?
<b>issues addressed</b>	the <b>issues addressed</b> - What were the evaluation questions? How were they decided? Who was involved in deciding the questions?
<b>collecting and analysing data</b>	the <b>methods employed</b> for collecting and analysing data - What were the principal techniques used? Who carried out the collection and analysis? What tools did they use? Are these techniques and tools recognised by the "evaluation community"?
<b>reporting the results</b>	the <b>timetable and arrangements for reporting the results</b> - What were the deadlines for reporting set out in the terms of reference? Were reports delivered on time? Was there a dissemination plan? What dissemination activities were actually carried out?
<b>quality control and assessment</b>	the arrangements in place for <b>quality control and assessment?</b> - Who was responsible for quality control? Who carried out the quality assessment? What criteria were used?
<b>follow up of results</b>	the <b>follow up of results</b> - Did those responsible for managing the activity respond to the findings? Was an action plan drawn up? Was its implementation monitored?

## RISK TO THE QUALITY OF THE EVALUATION

Identifying the risks to the quality of an evaluation is about analysing its context and management. The following sub-sections provide, for each of the main qualities of an evaluation (i.e. usefulness, coherence, robustness, impartiality, clarity, cost-effectiveness), a list of the main types of risk associated.

- Usefulness**
- Questions addressed not relevant to the purpose or intended users e.g. issues addressed are too detailed to be of interest to higher level decision makers (e.g. Parliament).
  - Information not available on time e.g. not finalised until after the decision to which it was supposed to contribute.
  - Inappropriate or no communication e.g. report not translated, no active communication to the relevant committees of the Parliament and the Council, unjustified confidentiality, communication restricted to making the document available on the intranet / internet.
- Coherence**
- Evaluation questions do not reflect the intended purpose; e.g. the purpose of the evaluation is to report on the extent to which objectives have been met for the purposes of accountability but the evaluation questions relate to testing the intervention logic (i.e. identifying cause and effect relationships).
  - The data set collected does not correspond to the questions; e.g. the objective of the evaluation was to judge the effects of training for young people who are socially excluded from the workplace, however the data collected on which the judgement was based came from the providers rather than the recipients of training.
  - Mismatch between data collected and analysis undertaken; e.g. statistical analysis carried out on the wrong type of data.
  - Incorrect "findings", due to inappropriate criteria or standards used.
  - Conclusions drawn do not follow from the data collected and the analysis carried out; e.g. conclusions overly definite given their basis, for instance, by presenting results as being conclusive when they could only be indicative given the nature or size of samples used.
  - Recommendations made do not correspond to weaknesses or problems identified and explained in the evaluation report.
  - The arrangements for managing the evaluation do not fit with its intended purpose, in particular, as regards the identities of the evaluator and project manager or the composition and use of a steering group; e.g. the evaluation is primarily for accountability purposes but is carried out internally or managed by services responsible for implementing the activity being evaluated.
- Robustness**
- There are many ways in which the results of data collection and analysis might fail to correctly reflect the underlying reality and thus render the evidence unreliable. These are largely related to the inherent limits associated with any given method. The following is a list of common examples:
- 'leading' questions asked in interviews;
  - questionnaires not filled in correctly e.g. due to inherent ambiguities or biases or self reporting;
  - official statistics for different regions or countries not being prepared on the same basis and therefore not being comparable;
  - 'sampling risk' i.e. the risk that a sample is not representative of the population from which it is drawn;
  - data deliberately manipulated or invented by the evaluator.
- Impartiality**
- Economic dependence of the evaluator on the evaluatees.
  - Inappropriate control over the evaluation process by one stakeholder group e.g. beneficiaries of the intervention evaluated.
  - Political pressure either exerted through the hierarchy or leading to self censure (i.e. unpalatable messages are avoided by participants in the evaluation process).
- Clarity**
- Absence of an executive summary to the main report adapted to the needs of higher level decision-makers and/or an overly detailed or complicated main report.
  - Ambiguity i.e. results mean different things to different people - this could be deliberate in order to mask disagreements between stakeholders in the evaluation process.
  - Conflicting messages e.g. mutually exclusive conclusions or recommendations presented.
  - Missing information or inadequate description e.g. of context, data collected, analytical methods used.

## Cost effectiveness

- Same result could have been achieved using less costly sources of management information e.g. monitoring, internal studies, audits.
- Budget advertised / assigned to the evaluation was disproportionate to the demands of the study eventually agreed.
- Overly burdensome management structure used e.g. too many participants in the evaluation process.

## Common risk factors

The following is a list of common factors associated with one or more of the risks identified above:

**Ill-defined purposes or questions** will tend to undermine the coherence, clarity and usefulness of the evaluation as it is difficult to a) develop appropriate data collection and analysis methods if the questions they are supposed to answer are not clear, and b) meet user needs that have not been sufficiently identified.

**A lack of pre-existing monitoring information** can limit usefulness and cost-effectiveness as information is costly to create through evaluations, thus reducing the scope for analysis and the number of questions that can be addressed.

**A low budget** could result in problems of coherence or robustness as cost pressures on evaluators may result in methodological short cuts (e.g. smaller sample sizes) or undermine the thoroughness with which data collection and analysis activities are carried out (e.g. use of less qualified or experienced researchers).

**An inexperienced project manager** or no single project manager could undermine almost all aspects of the evaluation.

**No tendering procedures or a poor response to calls for tender** should raise questions about the suitability of the evaluator in terms of their experience and/or impartiality. In addition, it might indicate cost-effectiveness issues e.g. few or no tenders received could be a sign that the contract value is too low to be profitable for experienced practitioners.

**A lengthy period between draft and final report or many drafts produced** could indicate general problems with the quality of evaluation. Alternatively, impartiality may have been compromised after initial messages proved unpalatable to some participants and the evaluator was put under pressure to change his judgements.

**Disagreements within the steering group** about the approach, conclusions or quality of the report could indicate, in particular, issues regarding the coherence and robustness of the work of the evaluator. Equally, they could indicate unpalatable messages and risks to impartiality.

**The involvement of external evaluator in subsequent or previous work in the same area** with the Commission raises issues of conflicts of interest which could call into question the impartiality of the evaluation. The same risk to independence applies to internal evaluators.

**The lack of a steering group** could present risks to the usefulness, coherence, robustness and impartiality of the evaluation. Steering groups are particularly relevant to ensuring that the needs of potential users of the evaluation are taken into account when the questions are decided. They also have a role to play in facilitating the evaluators access to information and providing methodological support and assessing quality. Furthermore, including a range of stakeholders in the activity / intervention in a steering group can help ensure that the judgement of its effectiveness is 'fair' or 'balanced'.

**A failure to meet deadlines** can be the sign of a generally badly managed project and so could be indicative of a range of quality issues. However, there is a very specific risk in terms of usefulness where the primary purpose of the evaluation was to support a specific decision.

**Poor communication** throughout the process.

## SECTION 2: AUDIT TESTING

### GATHERING PRELIMINARY INFORMATION

	<p>Preliminary information gathering will be necessary in order to establish the context of the evaluation and how it is managed as the basis for carrying out an analysis of the risks to quality and drawing up an audit programme.</p>
<b>Evaluation function of the DG</b>	<p>The <b>evaluation function</b> of the Directorate General (DG) is likely to be the principal source of information about the context of an evaluation and should be able to identify the project manager on any given evaluation.</p>
<b>Evaluation project manager</b>	<p>The evaluation <b>project manager</b> is the primary source for all the documents relating to the evaluation process; in practice, the project manager is very often from the evaluation function.</p>
<b>Websites</b>	<p>The Commission's <a href="#">Secretariat-General</a> is responsible for the evaluation of the European Union policies and initiatives. From its website (Smart regulation) there is a link to the evaluation activities of each DG.</p>

### AUDIT TECHNIQUES

	<p>Evaluations are exercises involving experts in the production of management information, and so, all the normal audit techniques associated with the audit of management information and the work of experts apply.</p>
<b>Interviews with evaluation stakeholders</b>	<p><b>Interviews with evaluation stakeholders</b> e.g. the evaluator, the evaluation project manager, the evaluation function, members of the operational unit responsible for activities evaluated and representatives of the primary users of the report are an essential source of information. They will provide much of the information necessary for establishing the context of the evaluation and how the process was managed. In addition, interviews can provide direct evidence about quality e.g. interviews with users will provide direct evidence about the usefulness of the evaluation. It will often be important to triangulate the views of different stakeholders to the evaluation as perceptions may vary amongst them about its quality.</p>
<b>Reviews of documents</b>	<p><b>Reviews of documents</b> produced by the evaluation and in response to its results e.g. evaluation plans, evaluation mandate, terms of reference, inception, interim, draft and final reports, quality assessments, responses to findings by operational units, action plans, follow up reports, meeting notes of the steering groups, meeting notes of institutions where the report was discussed (e.g. committees of the Council or Parliament, Committee of the Regions, the Economic and Social Committee), press reaction are all potential sources of information. These documents will provide much of the evidence necessary for establishing the context of the evaluation and how it was managed and can be used to corroborate information from interviews.</p> <p>Quality assessments will often be useful for highlighting risk areas. The final evaluation report itself will be the primary source of evidence about clarity and the coherence of the evaluation. However, much of the audit evidence should come from comparing documents e.g. the purpose of the evaluation as described in the plan to the questions in the inception report (coherence, usefulness) or the successive drafts of the final report (impartiality).</p>
<b>Observation</b>	<p><b>Observation</b> - in some cases it may be possible to attend evaluation related meetings as an observer e.g. attending follow up meetings could provide evidence about the usefulness of conclusions and recommendations.</p>
<b>Re-performance of analysis</b>	<p><b>Re-performance of analysis</b> using underlying data e.g. statistical analysis could be re-run using the same or equivalent software. In some cases, this will simply involve checking that the information in the main report accurately reflects that presented in annex.</p>
<b>Corroboration of underlying data</b>	<p><b>Corroboration of underlying data</b> e.g. check encoded data to underlying records such as completed questionnaires, circularisation of sample of respondents to check their existence, understanding and assessment of questionnaires used, compare results to similar studies carried out and published in the academic literature; checking publicly available information used e.g. official statistics, references to academic studies.</p>
<b>Expert opinion</b>	<p>An <b>expert opinion</b> on the final report could be used to help assess its coherence, robustness, cost-effectiveness, impartiality or likely usefulness e.g. the expert/s could be asked to comment upon the feasibility of recommendations, the suitability of methodologies used, cost-effectiveness of the study design, underlying assumptions made.</p>
<b>Focus groups</b>	<p><b>Focus groups</b> of users of the evaluation could be used to address the issue of its usefulness.</p>



## AUDIT GUIDELINES ON EVALUATION - APPENDIX B

# KEY ELEMENTS FOR BUILDING UP A PROGRAMME FOR AUDITING AN EVALUATION SYSTEM

### TABLE OF CONTENTS

#### INTRODUCTION

#### SECTION 1: Documenting the context and the system

Context  
Evaluation System

#### SECTION 2: Assessing the design of the system

Managing demand  
Ensuring the quality of the supply  
Encouraging the use of results  
Audit testing

### INTRODUCTION

This note provides practical guidelines on how to assess the quality of a European Commission Directorate General's (DG's) evaluation system. As there is considerable variation between the policy areas in terms of the Commission's responsibility for developing and implementing policy, there is consequently considerable variation in the numbers, types and uses made of evaluations. As a result, it is not practicable to develop a "one size fits all" audit programme.

However, a number of qualities of good evaluation systems have been identified in the literature on evaluation. In addition, the Commission has developed evaluation standards that are mandatory for all DGs (whose activities have addressees external to the Commission) and which provide the framework within which DGs must develop their evaluation systems. The suggested approach outlined below draws on both.

The main steps in the suggested approach are:

1 - document the context provided by the policy area and the evaluation arrangements in place;

2 - assess the design and test the operation of the system in terms of its ability to:

- manage the demand for evaluation;
- ensure the quality of evaluations carried out;
- encourage the use of results.

**Section 1** in this note covers documenting the context and the evaluation systems.

**Section 2** sets out a list of criteria that could be used for assessing the design of the system and outlines the main audit techniques that are mostly likely to be appropriate for testing its operation. Assessing the quality of individual evaluations is covered by Appendix A (Key elements for building up a programme for auditing an evaluation).

#### List of documents related

#### Audit Guidelines on Evaluation

#### Appendix A: Key elements for building up a programme for auditing an evaluation

#### Who to contact

If you feel that the information provided in this document could be improved, please do not hesitate to communicate your suggestions: [ECA-AMS.CONTACT@eca.europa.eu](mailto:ECA-AMS.CONTACT@eca.europa.eu).

## SECTION 1: DOCUMENTING THE CONTEXT AND THE SYSTEM

### CONTEXT

Evaluation needs can vary widely between policy areas depending on the nature of the activities carried out. Evaluation systems should therefore be adapted to their context taking due account of the costs involved.

Establishing the context involves gaining a sufficient understanding of:

- the legal framework in the policy area;
- the overall objectives in the policy area and the objectives of the individual interventions that fall within the remit of the evaluation system;
- the types of activity carried out, e.g. expenditure programmes, regulatory measures, financial instruments;
- the intervention logic underlying the way the policy is implemented;
- the resources available and the methods for managing them;
- the responsibilities of the different actors for delivering the different elements of the policy (e.g. DGs, Member States, regional or local authorities and partner organisations).

### EVALUATION SYSTEM

Documenting evaluation systems is about establishing the procedures and arrangements in place for:

- **programming and monitoring evaluations**, i.e. anticipating decision making needs, setting priorities, allocating resources and monitoring that plans are implemented;
- **ensuring quality in design, execution and reporting**, i.e. structuring the evaluation, coordinating participants, selecting evaluators, and controlling and assessing quality;
- disseminating and using results, i.e. deciding what should be disseminated and how, responding to findings and recommendations, establishing action plans, and following up action taken;
- **supporting evaluations**, e.g. providing resources, training, guidance, advice.

## SECTION 2: ASSESSING THE DESIGN OF THE SYSTEM

A good evaluation system should have measures in place to manage issues related to the **demand, quality of supply and use** of the evaluations. Listed below are a number of key qualities that an evaluation system should strive to achieve in order to ensure that it produces and uses sufficient quantities of good quality evaluations. For each of the qualities under the three headings, a number of indicators / good practices are highlighted that, if present, could provide the basis for a positive assessment.

### MANAGING DEMAND

#### Support from the top

This is about assessing the importance attached to evaluation within the DG. A good evaluation system should ensure that:

- High level decision makers value evaluation e.g. evaluation reports are read and commented upon, active involvement in the development of evaluation policy and plans, an evaluation mandate exists and was signed by the DG;
- Evaluation Function has a strong position in the DG, e.g. it is set up as a dedicated unit with a sufficient number of senior, experienced and permanent staff, it is in existence for a number of years, it reports directly to the Director General, it has budgetary control over evaluations/methodological work;
- Evaluations are routinely carried out and widely used in the DG.



**Reasonable expectations** Managing expectations is important for ensuring evaluation is supported and used. A good evaluation system should ensure that:

- Evaluation training is given to senior staff and/or those managing activities evaluated, covering the potential benefits and inherent limitations of evaluations;
- Stakeholders participate in the design and management of evaluations e.g. through steering groups;
- Awareness of users is raised and their needs are identified, e.g. through surveys;
- Feedback on the use of evaluation results is organised, e.g. internal publicity through seminars, workshops and the intranet;
- An active internal evaluation network, involving representatives of all operational units and those responsible for strategic planning and programming (SPP), meets regularly.

**Sufficient links with decision making** Creating links with decision-making processes creates an expectation for information from evaluations. A good evaluation system should ensure that:

- Procedures for ensuring monitoring and evaluation arrangements are explained in proposals;
- Decision-making documents are available for comment by those most likely to be aware of relevant evaluation results e.g. the evaluation function;
- Officials responsible for designing interventions are involved in their evaluation.

**A mix of incentives** A balance of "sticks, carrots and sermons" is important for ensuring sufficient evaluations are carried out. A good evaluation system should ensure that:

- Clear and realistic requirements to evaluate (sticks);
- Appropriate funding arrangements i.e. in terms of amounts available and who controls them (carrots);
- Evaluation is encouraged through awareness raising exercises such as seminars, workshops, training, the intranet (sermons).

## ENSURING THE QUALITY OF THE SUPPLY

**Training and support** Having the capacity to produce sufficient quantities of good quality evaluation efficiently requires a number of factors to be considered.

Training and support are crucial for ensuring the competence of those involved in the evaluation process and for learning over time. A good evaluation system should ensure that:

- Needs of staff in their respective roles in the evaluation process have been identified and tailored training / guidance developed e.g. for project managers, steering group participants, and programme managers;
- Evaluation training has been developed by the DG specific to its needs;
- Workshops / seminars are regularly held to exchange good practice and experience;
- Evaluation training included in induction training for officials of the DG;
- Helpdesk has been set up for advice on evaluation issues.

**Monitoring and audit systems** Making effective use of available information is important for ensuring the cost-effectiveness of evaluations. Periodic audit can help improve the performance of the system as a whole. A good evaluation system should ensure that:

- Evaluators have access to monitoring data for the activities evaluated;
- Monitoring systems are designed to facilitate evaluation;
- Evaluations are designed to facilitate meta-analysis or meta-evaluation<sup>1</sup>, e.g. to draw conclusions at a higher level or of more general relevance than would be possible on the basis of an individual evaluation;
- The evaluation system is periodically audited by the Internal Audit Capability of the DG.

**Planning evaluations** Planning is essential for ensuring the relevance and timely delivery of evaluation results. A good evaluation system should ensure that:

<sup>1</sup> Synthesis based on a series of evaluations. The term is also used to refer to an evaluation of another evaluation or of a series of evaluations.

- Mechanisms are in place for anticipating decisions that should or could be supported by evaluation results;.
- Realistic deadlines are set (e.g. taking proper account of the time it takes to procure external expertise);
- The purpose of evaluations and the reason for their selection are clear;
- Criteria have been developed for prioritising / selecting evaluations;
- Sufficient resources are available and allocated to implement plans;
- The implementation of the evaluation plan is monitored and reported on to senior management.

**Involving stakeholders where appropriate**

Stakeholders include both those involved in or affected by the activity being evaluated and stakeholders in the evaluation itself. A good evaluation system should ensure that:

- Main categories of stakeholders are identified;
- Consultation with stakeholders is valued and encouraged;
- A policy, requirements or guidance exists covering the composition and role of steering groups;
- Due consideration is given to involving those with access to data needed, those responsible for managing the activity evaluated, addressees and beneficiaries, and other intended users.

**Ensuring methodological quality**

Quality control during the course of the evaluation and quality assessments of final reports provide information and assurance to users about the reliability of the results. A good evaluation system should ensure that:

- Quality criteria specific to the DG / policy area have been developed;
- Quality requirements are included in the terms of reference for evaluators;
- The quality of successive documents produced by evaluators is controlled before they are accepted;
- The final evaluation report is subject to a formal quality assessment in writing;
- There are appropriate arrangements for ensuring the objectivity of published quality assessments.

**Choosing appropriate evaluators**

Much of the quality of an evaluation depends on the skills and knowledge of the evaluator. A good evaluation system should ensure that:

- Internal evaluators have sufficient experience of carrying out evaluations and/or have received appropriate training;
- Selection criteria for choosing external experts take account of experience of both evaluation and the policy area;
- Tender committees are made up of people with knowledge or experience of evaluation;
- Calls for tender and longer term evaluation needs and resources available are suitably publicised within the evaluation profession;
- Evaluators are required to disclose circumstances that might lead to conflicts of interest.

<b>ENCOURAGING THE USE OF RESULTS</b>
---------------------------------------

For evaluations to be used it is not sufficient just to create demand for them and ensure that evaluations carried out are of good quality. The evaluation system needs specific measures to ensure the results of evaluations are used both individually and in aggregate.

**Identifying users and their needs**

Evaluations should focus on the information needs of specific groups or individuals. A good evaluation system should ensure that:

- Consideration is given to the needs of users (e.g. internal and external, at the different levels in the hierarchy, other institutions) at the programming, planning and dissemination phases;
- There is coherence in the identification and treatment of user needs at the programming, planning and dissemination phases.

## **Ensuring the relevance of questions**

As evaluations usually involve considerable participation by those responsible for the management of activities evaluated, a challenge is usually to ensure that the needs of other users (particularly higher-level users, e.g. the Director General, the Commission, specific committees of Parliament and Council) are properly reflected when the evaluation questions are determined. A good evaluation system should ensure that:

- A policy, requirements or guidance provides for including standard questions in evaluations - this is particularly important for enabling conclusions to be drawn across evaluations that are relevant to the setting of political priorities and the allocation of resources;
- Users participate in the setting of questions, (e.g. involvement of Budget DG in steering groups encourages focus on efficiency and resource allocation issues);
- Incentives exist for addressing interesting / controversial issues ("sticks, carrots or sermons");
- Involvement of evaluation functions or the hierarchy in the processes for determining the questions to be addressed is encouraged;
- High level approval (e.g. by the Director General) is required at the stage at which questions are developed.

## **Making judgements and recommendations**

Evaluations need to produce interesting and credible results with recommendations that are relevant and feasible in order to maximise their impact and use. Self-censure and "blocking" by participants in the evaluation process can lead to controversial issues being avoided or the dilution of messages. A good evaluation system should ensure that:

- Evaluators are encouraged to make explicit judgements and specific recommendations (e.g. through terms of reference);
- There is an option for preserving confidentiality for the results of some types of evaluation.

## **Communicating findings**

Dissemination activities need to be properly planned and managed. A good evaluation system should ensure that:

- Requirements or guidance have been developed for disseminating the results of evaluations;
- Dissemination plans are drawn up at the inception stage or before;
- Dissemination activities are monitored to ensure results are successfully transmitted and to provide feedback on the impact of the dissemination methods used;
- Forms of dissemination are adapted to the needs of intended users e.g. meetings, press conferences, confidential briefings, publication on the internet;
- General publication is the norm rather than the exception.

## **Delivering evaluations in time**

Deadlines need to be set and managed with lessons learnt over time. A good evaluation system should ensure that:

- Project managers are aware of all the deadlines associated with the use of the evaluation results by the different users;
- Project managers report regularly to senior staff about the timing of ongoing evaluations;
- Where delays are unavoidable, they are anticipated, the consequences are thought through and appropriate action taken e.g. users notified, early findings are disseminated, the evaluation is scaled down or halted and resources allocated to projects that are more useful;
- Systematic delays are identified and corrective action taken e.g. to simplify procedures

## **Monitoring use and follow up of results**

Monitoring the use and following up the results of evaluations helps to provide evidence of the difference evaluations can make to the quality of decision-making and accountability. Thus, the value attached to evaluations will increase and hence they will be more widely used. A good evaluation system should ensure that:

- Managers of activities evaluated are required to reply formally to evaluation reports;
- Action plans are drawn up in response to evaluations and endorsed at a high level with their implementation monitored;
- The use of findings, results, conclusions and recommendations is monitored (e.g. by the evaluation function) to provide feedback on the usefulness of evaluations.

## AUDIT TESTING

### Documenting the policy context and the evaluation system

Most DGs provide summaries of their policy areas on their internet.

The evaluation function is likely to be the principal source of information about the evaluation system of a DG as it has the overall responsibility for coordinating the DG's evaluation activities. Information specifically about the evaluation system and recent evaluation reports will also be available on the internet sites of the DG.

The Commission's [Secretariat-General](#) is responsible for the evaluation of the European Union policies and initiatives. From its website (Smart regulation) there is a link to the evaluation activities of each DG.

### Audit testing

The normal principles and techniques apply when auditing evaluation systems.

### Interviews

- **Interviews with the main actors** e.g. the head of the evaluation unit, officials in the evaluation unit, evaluation project managers, members of the operational unit responsible for activities evaluated, users of evaluation reports (including members of the hierarchy). These will provide much of the information necessary for establishing the policy context and the evaluation system. In addition, interviews can provide direct evidence about the value attached to evaluation, its use, user needs, and evaluation experience of the different actors. Officials from the Internal Audit Capability may also have considerable information about the evaluation system if it has been the subject of an audit.

### Review of documents

- **Reviews of documents** are likely to be the main source of evidence about both the design and operation of the evaluation system e.g. evaluation plans, the evaluation mandate, evaluation guides, the documents relevant to the management and follow up of individual evaluations (terms of reference, inception, interim, draft and final reports, meeting notes of the steering groups, quality assessments, dissemination plans, responses to findings by operational units, action plans, follow up reports) as well as the main documents prepared by the DG or Budget DG in the framework of ABM/ABB (Annual Management Plans, Annual Activity Reports, Activity Statements, Annual Evaluation Reviews, Evaluation Highlights).

### Observation

- **Observation** - In some cases, it may be possible to attend evaluation related meetings as an observer e.g. steering committee meetings, workshops, seminars, meetings of the internal evaluation network.

### Expert opinion

- An **expert opinion** could be helpful in assessing the design of the evaluation system. In designing or developing their evaluation systems, DGs frequently use outside expertise. An external expert could help provide an opinion that takes into account knowledge of other DGs, internal organisation, and national/regional administrations in Europe or beyond. However, as many experts will have worked for the Commission, or may want to in the future, care must be taken to avoid a conflict of interest or a lack of independence.