EUROPEAN COMMISSION
DIRECTORATE-GENERAL
REGIONAL AND URBAN POLICY

DIRECTORATE-GENERAL
EMPLOYMENT, SOCIAL AFFAIRS AND INCLUSION

The Programming Period 2014-2020

MONITORING AND EVALUATION OF EUROPEAN COHESION POLICY

European Regional Development Fund
European Social Fund
Cohesion Fund

Guidance Document on Evaluation Plans

Terms of Reference for Impact Evaluations
Guidance on Quality Management of External Evaluations

April 2014

Quotation is authorised as long as the source is acknowledged.

Contents

1. INTRODUCTION 3

2. ROLE OF THE EVALUATION PLAN 3

3. ELEMENTS OF THE EVALUATION PLAN 5

   3.1. Objectives, coverage, coordination 5

   3.2. Evaluation framework 5

   3.3. Planned evaluations 6

4. HOW TO PLAN IMPACT EVALUATIONS? 8

   4.1. What is required by the Regulation? 8

   4.2. Timing of impact evaluations 9

   4.3. Necessary data 10

5. ADOPTION, REVIEW AND FOLLOW UP OF THE EVALUATION PLAN 11

ANNEX 1 : TERMS OF REFERENCE FOR IMPACT EVALUATIONS

ANNEX 2 : GUIDANCE ON QUALITY MANAGEMENT OF EXTERNAL EVALUATIONS

1. INTRODUCTION

The regulation for the 2014-2020 European Structural and Investment Funds (ESIF)[1] represents a radical change for operational programmes. The Common Provisions Regulation (CPR) emphasises programme objectives, the logic of intervention to achieve the expected results and the evaluation of effectiveness and impacts. Furthermore it requires from Managing Authorities and the Commission annual reporting on outputs and results, including findings of evaluations where available. In order to strengthen the contribution of evaluations to the effectiveness of the programmes, the CPR makes it compulsory for Managing Authorities to design evaluation plans at the beginning of the programming period.

> ***An evaluation plan*** *shall be drawn up by the managing authority or Member State for one or more operational programmes. The evaluation plan shall be submitted to the monitoring committee* ***no later than one year after the adoption of the programme****.* (Article 114(1) CPR).

In previous programming periods, evaluations have tended to focus more on implementation issues than capturing the impacts. For 2014-2020, the CPR requires Managing Authorities to carry out evaluations which assess the effects of the ESIF programmes. This is an essential element of the strengthened results-focus of the policy and is reflected in Article 56(3) of the CPR. The evaluation plans are therefore strategic documents setting out how these evaluations will be organised in order to provide evidence on effects for policy making.

> *During the programming period, the managing authority shall ensure that evaluations, including evaluations to assess effectiveness, efficiency and impact, are carried out for each programme* ***on the basis of the evaluation plan*** *and that each evaluation is subject to appropriate follow-up in accordance with the Fund-specific rules.* ***At least once during the programming period, an evaluation shall assess how support from the ESI Funds has contributed to the objectives for each priority****. All evaluations shall be examined by the monitoring committee and sent to the Commission. (Article 56(3) CPR)*

This document provides guidance to Managing Authorities on the contents of evaluation plans.

2. ROLE OF THE EVALUATION PLAN

Evaluation should be an essential part of the life cycle of a programme. Evaluations should serve to improve the effectiveness and efficiency of programmes as well as to assess their effects. They are meant to increase knowledge of what works and what does not in order for decision makers to make timely decisions to support the implementation of programmes and to draw conclusions for policy making. To support this knowledge building, the Commission intends to collect evidence from evaluations covering the same policy fields across Member States: where a sufficient number of good quality evaluations cover the same fields, it will carry out meta-analyses and to the extent possible build a repository of evidence for policy making. The results of these analyses will be disseminated widely. The Commission will also encourage exchange of information between Member States on available evidence and also on fields lacking strong evidence.

The role of the evaluation plan is central to achieving this aim: it will support quality evaluations as well as their effective use by Managing Authorities; it will facilitate sharing of knowledge on what works and how in different policy fields; and, ultimately, it will contribute to the design and implementation of evidence based programmes and policies.

[1] Regulation (EU) No 1303/2013 of the European Parliament and the Council of 17 December 2013 laying down common provisions on the ESI Funds, referred to throughout this document as the CPR (Common Provisions Regulation).

Tasks of the evaluation plan
- Improve the quality of evaluations through proper planning, including through identification and collection of necessary data (Article 54(2) CPR);

- enable informed programme management and policy decisions on the basis of evaluation findings;

- provide a framework to plan impact evaluations (Article 56(3)CPR)

- ensure that evaluations provide inputs for annual implementation and progress reports;

- facilitate the synthesis of findings from different Member States by the Commission and the exchange of available evidence;

- ensure that resources for funding and managing the evaluations are appropriate.(Article 54(2) CPR).

In particular, the evaluation plan will ensure that evaluations are available on time to inform the different reports on results that Member State authorities will have to deliver from 2016 onwards: annual implementation reports and progress reports in 2017 and 2019 (Articles 50 and 52 CPR); as well as the summary report and synthesis of findings of the available evaluations of programmes to be transmitted by the Commission each year from 2016 to the Council, the European Parliament, the European Economic and Social Committee and the Committee of Regions (Article 53). The strategic reports of the Commission due in 2017 and 2019 will draw on Member States progress reports (Article 53 CPR).

Furthermore, at least twice during the programming period an evaluation should assess the effectiveness, efficiency and impact of the joint support by the ESF and the Youth Employment Initiative including for the implementation of the Youth Guarantee. The first should be completed by end 2015 and the second by end 2018 (Art. 19(6) ESF). The main findings of the evaluations assessing the joint support from the ESF and the YEI should also be presented in the annual implementation reports and progress reports (Art. 19(4) ESF).

The evaluation plan should also ensure that the essential material is available in time for the Managing Authorities to meet their obligation laid down in article 114.2 (CPR) to submit to the Commission a report summarising, for each operational programme, the findings of evaluations carried out during the programming period, by 31 December 2022 (Article 114(2) CPR).

Who?	What?	When?	Reference
Member State/Managing Authority	In the AIR: Synthesis of the findings of all evaluations of the programme (including YEI where appropriate)	By 31 May each year from 2016 until 2023, by 30 June in 2017 and 2019	CPR Art. 50
Commission	Summary report based on the annual implementation reports of the Member States; as well as a synthesis of the findings of the available evaluations of programmes	Each year from 2016	CPR Art. 53
Member State	In the progress reports: Progress made towards achievement of the Union strategy for smart, sustainable and inclusive growth, as well as of the Fund-specific missions (including YEI where appropriate)	By 31 August 2017 and by 31 August 2019	CPR Art. 52
Commission	Strategic report summarising the progress reports of the Member States	By 31 December 2017 and 31 December 2019	CPR Art. 53
Member State	Report summarising the findings of evaluations carried out during the programming period	By 31 December 2022	CPR Art. 114

3. ELEMENTS OF THE EVALUATION PLAN

The evaluation plan should contain three main parts as outlined below.

3.1. Objectives, coverage, coordination

- An introduction to the plan setting out its main objectives.
- Coverage and rationale: an explanation of which programmes and which ESI Funds are covered by the plan and why (Article 114(1) CPR). The evaluation plan may include evaluations of interventions from other programming periods when deemed useful for improving the 2014-2020 programming.
- An analysis of relevant evidence available in order to decide where the evaluation efforts should be most concentrated. This evidence could be found in evaluations carried out during the preceding periods, EU ex post evaluations, evaluation literature and preparatory studies for this programming period. When operational programmes are designed at regional level, the Commission suggests that this collection of evidence is coordinated at a national level as many interventions are similar in different programmes.
- It would also be good practice to set up mechanisms within a Member State for coordination and exchange between Managing Authorities on evaluations planned, evaluation findings and methodologies. This would allow for a better coordination of evaluations carried out in certain policy fields and favour exchange of knowledge and practices between managing authorities. A specific area of coordination is the integrated approach according to Article 15(2) of the CPR. For example, if an integrated approach is applied to address the specific needs of a target group at high risk of social exclusion through different operational programmes, the evaluations of the relevant interventions could be coordinated. The programmes' Technical Assistance could be used to support this process.

3.2. Evaluation framework

The evaluation plan should specify[2]:

- The **evaluation function** with a clearly defined responsibility for designing and delivering the evaluation plan, and coordinating, monitoring and promoting the quality of evaluation activities throughout the whole evaluation cycle.
- Description of the **evaluation process** led by the Managing Authority (responsibilities of involved bodies: evaluation steering group, technical working groups, scientific or other expert academic input, monitoring committee etc.).
- The **involvement of partners** in evaluation (Art. 5(2) and (3)(d) and art. 49(4)) within the framework of the monitoring committees or in specific working groups established by the monitoring committees; and their consultation on the report(s) summarising the findings of evaluations due by 31 December 2022.
- The source for **evaluation expertise** (internal/ external/ mixed) and provisions ensuring the functional independence of evaluators from the authorities responsible for programme implementation (Art. 54(3) CPR)[3].

[2] Some of these aspects may be covered by other documents: in this case the evaluation plan should explain which parts of the framework are set out in other documents, refer to those and provide their URL where available.

- Possibly a **training programme** (for example, seminars, workshops, self-study and working with other evaluators) for people from Managing Authority dealing with evaluation.
- A strategy to ensure **use and communication** of evaluations: how their findings will be followed up; how the evaluations will be made public (Article 54(4) CPR); how they will be transmitted to the Commission. Transmitting final evaluation reports through SFC[4] with the Terms of Reference, the budget, and the evaluation methodology is good practice; this will allow the Commission to provide examples of methodological approaches used in different fields to the evaluation community, to analyse the evidence produced and as far as possible build a repository of evidence for policy making.
- An overall **timetable** showing how the evaluations will feed into implementation and the various reports on programmes;
- The overall **budget** for implementation of the plan (covering the cost of evaluations, data collection, training etc.). Including a budget, human resources and possibly a training programme contributes to meeting the legal obligation of Member States to provide the resources necessary for carrying out evaluations (Article 54(2) CPR). If technical assistance is used for the purpose of the plan, the corresponding amount should be set aside in the TA budget.
- A **quality management strategy** for the evaluation process: drafting good terms of reference and managing contracts have an important role to play in delivering good evaluation results. See guidance in Annexes 1 and 2.

3.3. Planned evaluations

The plan should contain a list and timetable of the evaluations to be carried out throughout the period, including an explanation for the selection of the themes covered. If the plan covers more than one operational programme or more than one fund, it is important that a balance of themes among the programmes and funds is ensured.

It is clear that new evaluation needs may emerge in the course of the programme life cycle. The list thus needs to be regularly reviewed by the Monitoring Committee. It is suggested that the theme, scope and broad focus of the methodology of evaluations planned beyond a 3 year period are indicative. More detailed evaluation questions can be developed nearer the time of launching the evaluations.

On the other hand, it is important to plan sufficiently in detail the impact evaluations required by the regulation for each priority axis at an early stage. This would set out the practical arrangements for the organisation of impact evaluations (number of contracts, timing and coverage) and ensure that appropriate data will be available (see below). For example, if several regional OPs fund similar priority axes and within them have similar specific objectives and interventions, it could make sense to plan one common evaluation to evaluate the impact of these priority axes.

The evaluation plan is not restrictive: Managing Authorities may also conduct *ad hoc* evaluations if needed during the course of the programme life cycle.

In addition to the theme or topic, the evaluation plan should specify for each evaluation:

---

[3] See section 2.4. on functional independence in the Guidance document on ex ante evaluation http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#1

[4] Structural Funds Common database

- **Subject and rationale**, including the background, the coverage, the main approach (process or impact evaluation) and the main guiding evaluation questions. These can be adapted and further specified in the ToR of the evaluations.
- **Methods** to be used and their **data requirements**: according to the evaluation subject, different methods may apply: a process evaluation may use data analysis, interviews, surveys while an evaluation on the effects of the programme may involve other methods such as literature review, focus groups, case studies or setting up comparison groups (see below).
- Commitments that particular **data sets** required for certain evaluations will be available or will be collected and the timeframe; experience shows that the lack of systematic collection of evaluation data significantly increases the cost of collecting them retrospectively (the evaluator has to reconstruct ex post a set of data) or leads to using less rigorous methods (where collecting these data is not feasible). This requirement for availability of data is part of Member States obligation imposed by Article 54(2) CPR and by the general ex ante conditionality on statistical systems and result indicators.

Public data registers

Particular attention should be given to public data registers storing time series data covering samples or total populations of entities and participants. In order to use such data, contacts with authorities holding registers should be made sufficiently in advance. If some particular data are needed, the managing authority may need to negotiate with the relevant authority (e.g. national statistical office) to collect these additional data through their questionnaires. These interviews would have to be scheduled as soon as possible, when drafting the evaluation plan.

- **Duration and a tentative date**. These are linked with the evaluation subject and coverage and the methods selected, e.g. an impact evaluation can only be carried out once results are achieved, while some methods, where the necessary data have to be collected, are more time consuming than others.
- **Estimated budget** for each evaluation: the cost is linked to the selected methods and the duration of the contract.

4. HOW TO PLAN IMPACT EVALUATIONS?

Article 56(3) requires that, at least once during the programming period, an evaluation shall assess how support from the ESI Funds has contributed to the objectives for each priority axis. Article 50(4) CPR requires that the annual implementation report submitted in 2017 assesses progress towards achieving the objectives of the programme, "*including the contribution of the European structural and investment funds to changes in the value of result indicators, when evidence is available from relevant evaluations*."

In the 2014-2020 programming period, both result orientation and thematic concentration make it necessary to design programmes focussing their resources on a few objectives in order to maximise their impact. Their expected results shall be measured with result indicators and the programme effects assessed with impact evaluations. **In other words, the planning of impact evaluations must be based on sufficiently focussed programmes with a limited number of objectives clearly specifying the intended change.** This will help to limit the number of impact evaluations and to ensure that evaluation efforts are proportionate to the available funding.

4.1. What is required by the Regulation?

Results are at the core of the ESIF programmes. They are routed in the intervention logic of the programmes and some of their measurable dimensions are captured by result indicators reflecting the main outcomes expected for each specific objective. The contribution of the programme to these results must be assessed by impact evaluations covering each priority axis. This impact needs also to be assessed *"in relation to the targets under the Union strategy for smart, sustainable and inclusive growth and, having regard to the size of the programme, in relation to GDP and unemployment in the programme area concerned, where appropriate"* (Article 54(1)). The evaluation plan should specify how deeply the programme's contribution to the observed results (its impact) will be evaluated. This will depend on the nature of the result, the available knowledge in the field of intervention (if evaluations were already carried out on the same kind of intervention and/or the impact is likely to be achieved in the longer term, a meta-analysis of the findings of such studies or evaluations would suffice) and the importance of understanding the impact of some interventions for policy making. Thus the Regulation does not require every intervention to be evaluated, nor the same evaluative approach to be applied for all results across the programme. Each specific objective should be covered, however not to the same extent.

The evaluation plan should explain how the evaluation process will provide evidence to allow for overall conclusions on the contribution of each programme priority axis to their objectives. The evaluations planned should also allow the Managing Authority to conclude on the impact of the programme in relation to the targets of the Union strategy for smart, sustainable and inclusive growth. The evaluation plan is thus a strategic document which will accompany the programme throughout its life and support its result orientation.

Concluding on the impact of programmes or priorities in relation to the targets of the Union strategy for smart, sustainable and inclusive growth will not in most cases be possible using quantitative analysis, due to the amount of resources allocated to the programmes as compared to the needs and the contribution of other factors. Thus, in most cases, the assessment will be qualitative and take into account all the factors contributing to the evolution of these targets.

4.2. Timing of impact evaluations

As a general rule, impact evaluations should be scheduled as late as necessary to give time for the expected results to appear and as early as possible to allow their findings to feed into the policy process. Benefitting from the results of early impact evaluations may help to improve the interventions if they continue to be supported by the programme (or by other sources of funding).

Achieving this balance is a challenge. The planning has to be realistic and consider the nature of the interventions to be evaluated: for example, the impact of support to individuals or SMEs can be assessed earlier than the impact of educational or large scale infrastructures or RDT projects.

When results materialise in a longer time scale (e.g. infrastructure projects), impact may only be assessed after the end of the programming period: a possibility to cope with this difficulty is to carry out evaluations of the impacts of similar interventions supported in the 2007-2013 programming period. This can make sense as for many interventions it takes some years before the effects are fully realised and would bring interesting findings to improve the programme implementation. Ex post Cost Benefit Analysis (CBA) is an evaluative approach that can be applied to large scale infrastructures only several years after it has become operational.

In relation to a training intervention targeted at the unemployed, the question is over what time scales higher rates of employment and wages might materialise. It is a well-established feature of training programmes that in the short run they tend to reduce employment among participants as they tend to divert unemployed trainees away from job search due to their attendance at courses. Thus, if impacts are calculated too soon they may well be negative. Thus, a compromise has to be reached between what is reasonable for a follow-up interval from the perspective of the intervention and the need of Managing Authorities for timely evidence on impacts.

The appropriate timing of different evaluations suggests that Managing Authorities will need to carry out impact evaluations at different points in time. They may also need to work with different evaluation teams offering specific skills in different policy fields and evaluation methods (specialists in innovation may not have a deep knowledge on environmental issues, and experts in counterfactual impact evaluations may not be experienced in theory based evaluations).

Where possible, impact evaluations should be carried out in time to feed the report on evaluation results due by 31 December 2022. The main purpose of this report is to support the ex post evaluation that is under the responsibility of the European Commission in close cooperation with the Member States.

4.3. Necessary data

Evaluations capturing the impacts of priorities, whatever the methodological approach selected, will require data on supported entities and individual participants (micro-data).

Data on supported entities are also required as information on support from the Funds (Annex XII of the CPR Regulation). They are important for the Managing Authorities to follow and understand how the programme progresses.

Micro data are needed in order to carry out impact evaluations of interventions targeted on individuals. The CPR stipulates in Art. 125(2)(d) the legal requirement that Managing Authorities shall establish a system to record and store in computerised form data on each operation, including data on individual participants in operations.

Some types of impact evaluations will also need data coming from sources of information other than monitoring systems (e.g. social security, unemployment records or tax records). When planning their impact evaluations, Managing Authorities need to take into account possible data protection issues, especially in relation to sensitive data on participants.[5]

Two broad categories of impact evaluations are widely recognised:

- Theory-based impact evaluation, which follows each step of the intervention logic identifying causal links and mechanisms of change, answering the questions *why* and *how* an intervention works. This approach mainly produces a qualitative estimate of the impacts. It is based on approaches such as Process tracing, Contribution Analysis, General elimination method and includes methods such as literature review, text analysis, interviews, surveys, focus groups and case studies.
- Counterfactual impact evaluation, which uses control or comparison groups. This method is useful in answering *how much of the change* is due to the intervention and *for whom*[6].

  Counterfactual impact evaluation requires:

  1. a credible control or comparison group,
  2. a large enough number of participants and entities and controls for statistical significance, and

[5] Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data provides definitions of personal and sensitive data.

[6] In any case, the findings of counterfactual evaluations need to be complemented with qualitative methods to understand the mechanisms leading to the impact (or lack of impact) when it comes to policy decisions concerning the intervention.

3. good data on supported and non-supported participants and entities to compare results. It will be far easier to meet all of these conditions with appropriate planning from the outset, hence the importance of a good evaluation plan.

The two methodological approaches will include the analysis of the effects on participants and entities as an input to the judgement on the impact of the programme: the counterfactual impact evaluation on, say, an aid scheme for enterprises will compare the situation of the supported enterprises with a comparison group (comprising enterprises who were targeted but were not subject to the intervention) both before and after the completion of the operation. Similarly, to evaluate a training scheme for unemployed people, those trained (treated) will be compared to a similar group of eligible unemployed non-trained people or alternatively to unemployed people to whom a different aid was provided. As the appropriate selection of control groups is a complex exercise, it is recommended that commissioning authorities take early steps to identify potential controls and associated data requirements.

Theory based evaluation will look at the results of the interventions and assess the mechanisms leading to them, and will complement this analysis with other methods to explore other factors or alternative hypotheses before concluding on the impact of the programme on the observed result.

For further information on theory-based and counterfactual methods, see the Evalsed sourcebook on methods and techniques at
http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm
and the Practical Guidance for ESF Managing Authorities on the Design and Commissioning of Counterfactual Impact Evaluations at
http://ec.europa.eu/social/BlobServlet?docId=10805&langId=en

5. ADOPTION, REVIEW AND FOLLOW UP OF THE EVALUATION PLAN

The evaluation plan shall be submitted to the monitoring committee no later than a year after the adoption of the programme (Article 114(1) CPR). Due to the strategic dimension of the evaluation plans and to ease exchange between managing authorities, the Commission recommends that they are made public as are the programme (Article 115 (1)(b) CPR) and the AIR, for example by posting them on the website of the managing Authority.

Among the functions of the monitoring committee are the examination and approval of the evaluation plan and its review (Article 110(2)(c) CPR). The Commission recommends that the monitoring committee reviews the implementation of the evaluation plan at least once a year; it must approve any amendments which are deemed necessary. Monitoring committees can also set up specific working groups to involve relevant partners in evaluation.

The choice of the interventions, the impact of which will be evaluated more in depth is a strategic one: approving and reviewing the evaluation plan should be preceded by a debate in the monitoring committee on the expected main results of the programme, their policy importance, their timing and the availability of evidence in the field.

The monitoring committee will also examine the progress in the implementation of the evaluation plan and the follow up given to the findings of evaluations (Article 110(1)(b)). This shall be reported within each annual implementation report (Article 111(4)(a) CPR).

The Commission will analyse the content of the evaluation plans and the findings of the different evaluations when they are available: this will feed into its reporting to the other European Institutions as well as its activities directed at increasing the knowledge base on ESIF effects and disseminating good evaluation practices amongst regions and Member States. To ease the process of collecting the evaluation plans, the Commission recommends sending them through SFC.

ANNEX 1

TERMS OF REFERENCE FOR IMPACT EVALUATIONS

➢ INTRODUCTION

This Guidance is intended to assist national and regional authorities in the preparation of Terms of Reference (ToR) for impact evaluation with the main objective of improving the quality of them within Cohesion Policy programmes. The guidance is primarily meant to apply for evaluations to be carried out by external consultants but can provide also some advice for evaluations done within a national or regional administration.

For the 2014-2020 period, there are new requirements for impact evaluations to be carried out. The Regulation (EU) 1303/2013 on common provisions on the ERDF, the ESF, the Cohesion Fund, the EAFRD and the EMFF (art. 54- 56) stipulates that:

"…Member States shall provide the resources necessary for carrying out evaluations, and shall ensure that procedures are in place to produce and collect the data necessary for evaluations, including data related to common and where appropriate programme-specific indicators….

Evaluations shall be carried out by internal or external experts that are functionally independent of the authorities responsible for programme implementation. The Commission shall provide guidance on how to carry out evaluations…

During the programming period, the managing authority shall ensure that evaluations, including evaluations to assess effectiveness, efficiency and ***impact*** *are carried out, for each programme on the basis of the evaluation plan …. At least once during the programming period, an evaluation shall assess* ***how support from the ESI Funds has contributed to the objectives for each priority****. All evaluations shall be examined by the monitoring committee and sent to the Commission." (emphasis added)*

There are no specific requirements regarding ToR in the regulations governing Cohesion Policy. This guidance, based on the Commission's own experience and analysis of evaluations carried out in the Member States provides advice on the content and structure of the ToR in order to obtain good quality evaluations.

➢ DEFINITION OF TERMS OF REFERENCE

The Terms of Reference is a key document in the evaluation process, as it defines all aspects of how an evaluation will be conducted. It presents the objectives of the evaluation, the role and responsibilities of the evaluator and evaluation client and the resources available to conduct the evaluation.

ToR are developed during the planning phase of the evaluation process and are used to hire the evaluator on a competitive basis. Ensuring a high quality evaluation depends on how accurate and well-specified the ToR are.

Typically, the ToR provide the basis of the contractual arrangement between the public authorities who commission the evaluation and the evaluators responsible for carrying out the work. Depending on national legislation, there may be room for other procedures as well, where the commissioner of an evaluation first publishes a framework of ideas and based on the replies to this first call, a negotiated procedure can be put in place for the evaluation.

➢ WHAT TERMS OF REFERENCE SHOULD CONTAIN

The content and format for a ToR will vary based on legal and administrative requirements, local practices and the type of tasks. The ToR may range from a prescriptive list of tasks and operations to a flexible document that indicates the general approach and methods to collect and analyse data.

However, ToR should be as concise as possible, giving the evaluator all the necessary information. A typical length is 5-10 pages, excluding the administrative annexes that may accompany the core ToR to provide additional information to facilitate the work of the evaluator. These annexes may be the list of documents to be consulted, data on the intervention to be evaluated, a code of conduct, etc.

The ToR should include as a minimum the sections below.

Background and context

3.1 Evaluation purpose and target audience

3.2 Evaluation objective and scope

3.3 Evaluation questions and tasks

3.4 Approach and Methodology

3.5 Timing and deliverables

3.6 Evaluation team composition and required competencies

3.7 Management arrangements

3.8 Budget and Payment

3.9 Proposal Submission

➢ Background and context

This section should provide information on the background of the programme, project or activity to be evaluated. The description should be concise (maximum one page) and focus only on the issues pertinent for the evaluation:

- The intended objectives of the intervention to be evaluated should be clearly stated as well its rationale and scale – the logic of the intervention;

- The timeframe and the progress achieved at the moment of the evaluation;
- Key stakeholders involved in the intervention (including main clients, implementing agencies, others);
- Key elements at international, national or regional level relevant for the evaluation, including organisational, social, political and economic factors which have an influence on the intervention's implementation;
- Any previous study or evaluation conducted on the intervention.

More detailed information (detail on the theory of change, logical framework, indicators, etc.) may be included or referenced in the annexes.

➢ **Evaluation purpose and target audience**

This section outlines why the evaluation is to be conducted and identifies the key users of the evaluation findings. In general, an evaluation is carried out to understand the mechanisms and the impact of an intervention and to improve intervention design or management. These elements should be considered when writing this section:

- The whys for conducting an evaluation (i.e. relevant changes in the programme or its environment, fixed term, etc.);
- What the evaluation seeks to accomplish;
- Who will use the evaluation results (key users and target audience);
- How the evaluation results will be used.

Being clear about these elements indicates to the evaluator how the evaluation is intended to fulfil its accountability purpose.

➢ **Evaluation objective and scope**

The *objective* of the evaluation reflects what the evaluation aims to find out. It can be to quantify impact and/or to analyse the mechanisms producing the impact.

There should be no more than two or three objectives. Generally, it is recommended to explore few issues deeply rather than examine a broader set superficially.

The *scope* delimits the focus of the evaluation. Details here could include the time period, the geographical and thematic coverage of the evaluation, the target groups and the issues to be considered. The scope of the evaluation must be realistic given the time and resources available.

➢ **Evaluation tasks and questions**

The commissioning authority should have an idea of the work it wants the evaluator to undertake, whether it is a literature review, an analysis of administrative data, carrying out surveys or case studies or statistical analysis. These should be structured into a series of evaluation tasks with a logical sequence in terms of building up evidence on the

subject being evaluated. Under each evaluation task there will be specific evaluation questions.

The evaluation questions flow from the objectives and tasks of the evaluation and should correspond to a real need for knowledge, understanding or identification of new solution. The conclusions of the evaluation must clearly answer these questions, based on the evidence presented and analysed, incorporating the evaluator's judgment. There is no single evaluation method that can answer all possible questions.

Some practices should be followed when deciding on the questions to pose:

- the issues of greatest concern should be addressed by specific evaluation questions;
- the questions should be answerable – the commissioner of the evaluation should have an idea of the data and methods which could be mobilised to answer the question as well as whether or not this is the appropriate time to launch the evaluation.

An impact evaluation could address the following questions[7]:

- What change can be observed in relation to the objectives of the intervention?
- To what extent can observed changes be attributed to the intervention?
- Are there unintended impacts?
- What mechanisms delivered the impact? What are key contextual features for these mechanisms?
- Does the impact vary by subgroup within the main target group?
- Will short-run effects of the intervention differ from those in the long run?

An evaluation will always have conclusions, which answer the evaluation questions. It is not always necessary for an evaluation to contain recommendations. The commissioning authority needs to reflect on the best process to turn evaluation conclusions into changes in practice and policy. This can be through evaluation recommendations, which should, however, be discussed by the evaluator with the commissioning authority and those responsible for implementation, in order that the evaluator can understand the feasibility of different solutions. This is part of the process of deciding the evaluation purpose and use.

[7] Remember that this paper focuses on impact evaluations. Evaluations during the programming period can also revisit the intervention logic and resource allocation of a programme; ask questions to improve the management or implementation of a programme; focus on creating a shared vision between stakeholders on policy needs or outcomes (participative approaches). The detailed evaluation questions and methods will differ from what is discussed here.

➢ Approach and Methodology

Under the section on *Approach and the Methodology* there are two possibilities:

- The commissioning authority can indicate a preference for methods to be used, for all tasks or for a part of them, in detail or only the general approach, or

- it can leave the methods open and ask the tenderer to propose the precise combination of methods to be mobilised in carrying out the evaluation (or if possible, plan a step in the call for tender process to discuss with the competing tenderers the methods proposed).

In both cases, the commissioning authority needs to have an understanding of the potential, limits and basic technical features of methods. This is necessary for the commissioning authority to judge the appropriateness of the methodology proposed and whether it will make it possible to answer the evaluation questions under the various tasks.

In both cases also the ToR need to give scope to the tenderer to demonstrate its expertise in the methods. Note that the higher the level of detail in methodology description within the ToR, the higher the risk the proposals will simply replicate what the ToR stated. This would make it difficult to judge the different evaluation proposals. To prevent this, in cases where the commissioning authority want to prescribe the method, the ToR could request further detail to be added in the tender documentation (e.g., description of how the methods will be combined; motivated suggestion of case studies; proposal on the method based on available data with a discussion of strengths and weaknesses of data and how these will be addressed, a discussion on the details of method to be used with a motivation, etc.).

Impact evaluations should always start with a review of the theory of change which underpins the intervention being evaluated, whether this is explicit or implicit in programming documents. This can be done through literature review, interviews and surveys. The ToR should ask for an outline of what data sources will be used for this review.

Theory-based evaluations should answer questions concerning how and why a programme has, or has not, had an impact. A theory-based approach will investigate the causal linkage that relates inputs, activities and outputs to impacts. Its main goal is to explain why a given change has occurred and how an intervention has caused that change. Thus, a theory-based approach will put under accurate scrutiny all the assumptions that underlie the causal linkage. The ToR will ask the tenderer to identify methods to be used to gather information on causal links, what actually happens on the ground, behavioural change, etc. These can include surveys, interviews, focus groups, case studies, etc.

If a commissioning authority decides to verify a causal relationship between its programme and observed changes and in estimating its contribution to the change, it might choose a *counterfactual* evaluation approach. In order to carry out such an evaluation correctly, this type of evaluation should also start with the theory of the intervention in order to identify the outcome variables to be examined. A ToR for a counterfactual evaluation should include information on data available to construct a comparison or control group. It may ask the tenderers to assess the quality of this data and its strengths and weaknesses and to propose which among various techniques will be used (e.g., which discontinuity design(s), which forms of matching, etc.).

- **Timing and deliverables**

This section of the ToR details all the deliverables and their deadlines for the evaluation contract.

The list of deliverables should also include details related to timing, format, content, length, intended audience and the expected review process. The language of the deliverables should be specified as well. An abstract in English is recommended as it will help to disseminate the evaluation results beyond national borders and contribute to the accumulation of evaluation knowledge across Europe. An indicative number of pages for each deliverable could also be provided to give a total estimate for the work to be carried out and to signal a need for concise, analytical reports, rather than long descriptive texts.

At the minimum, the deliverables should include:

- an *inception report*, the detailed description of the methodology to answer the evaluation questions as well as the proposed source of information and data collection procedure. The inception report should also indicate the detailed schedule for the tasks to be undergone (work plan), the activities to be implemented and the deliverables. The role and responsibilities of each member of the evaluation team should be stated as well.
- a *draft evaluation report*, to be discussed among the relevant stakeholders in order to provide comments;
- *final evaluation report*, including:
    - Executive summary
    - Intervention description
    - Evaluation purpose
    - Evaluation methodology
    - Findings
    - Conclusions (answers to the Evaluation Questions)
    - Recommendations (if required)

- Annexes (list of people interviewed, key documents consulted, data collection instruments)

For more complex evaluations, it is advisable to ask one or more interim reports as well as monthly progress reports. Generally, it is also recommendable to foresee a system of economic incentives/disincentives to assure the quality of the deliverables and their on time delivery (i.e., linking payments to acceptance of deliverables, see below).

- **Evaluation team composition and required competencies**

The ToR must specify the procurement process for selecting an evaluation team or individual evaluators. This section should outline the mix of knowledge, skills and experience needed to carry out the contract.

This should include prior experience in design and leading evaluation; data analysis skills; knowledge of the regional and institutional context; technical competence in a specific sector; process management skills; language proficiency.

In this section the type of evidence required to demonstrate the claims of knowledge should also be specified. It is a good practice to require the potential candidates to submit two/three work sample when responding to ToR.

In the case of a team, it is recommended to ask for an evaluation expert to be nominated as leader while the other team members could be specialists in their respective area. In complex evaluations, the role of a leader or co-ordinating expert can be the crucial difference between a good and a poor evaluation. The role and the responsibilities of all the members should be defined precisely.

The requirements for the independence of evaluation should be stated in the ToR. Putting in place management arrangements that will support the independence of those evaluators chosen and requesting confirmation that there are no conflicts of interest within the potential evaluator is important. Equally important is to ask the tenderers to detail their systems of quality assurance for each deliverable and to ensure that appropriate resources (both in time and quality) are allocated to this function.

- **Management arrangements**

This section will describe the governance and management arrangements for the evaluation. These should clarify expectations, eliminate ambiguities and facilitate the conduct of the evaluation.

This section of ToR outlines:

- Requirements in relation to the detail the tenders should provide on the specific role and responsibilities of each and every component of the evaluation team. This should include a breakdown of the days input by task, by team member.

- the specific role and responsibilities of the evaluation client, such as providing in a due time comments on all the deliverables, assess the evaluation team in all the steps of the evaluation, providing contact and information, etc.

- the participation of other relevant stakeholders; especially during evaluation which involve many organizations (ie. joint evaluation) it is necessary to clearly define the role and responsibilities of each stakeholder involved in the evaluation process.

- The involvement of external experts: it is good practice to ask experts in the field to give input and review the deliverables at different stage of the evaluation.

The management arrangement includes also the lines of processes and responsibility for approving the deliverables.

Other issues to be considered include logistical questions such as office space, equipment, materials, etc.

➢ **Budget and Payment**

In this section a total amount of financial resources available for the evaluation (consultant fees, travels, allowance, etc.) should be outlined. Flat rate approaches can be useful to concentrate the later contract management on questions of quality of content.

In case of a limited budget, it is a good practice to suggest an indicative budget and then to leave those competing for an evaluation by open tender to suggest what they would be able to provide for the budget available. This allows value-for-money assessments to be made. An alternative is to leave space to competitors to propose their estimates based on the tasks seen as necessary. Generally, a breakdown of costs by tasks should be encouraged (eg. data collection, report preparation, fieldwork, etc.) in order to facilitate the proposals comparison. The ToR should also ask for the cost per day of different team members in order to understand the allocation of the work among senior and junior evaluators.

This section often includes any pertinent details related to payment. It is strongly recommended to connect payment to deliverables (e.g., linking the payment to the approval of the inception report, the draft of evaluation report and the final report).

➢ **Proposal Submission**

ToR are typically used to request proposals from bidders as part of a public and competitive tenders. In this case, the ToR should detail the instructions concerning format, content and submission of the proposals.

The details should include:

- Structure (indicating the items to include);

- Deadline;

- Criteria and timeline for proposal judgment;
- Opportunities for clarification, indicating the modalities and deadlines for questions.

➢ FINAL COMMENTS

This guidance does not cover exclusion, selection and award criteria and how these will be combined to select the contractor, as these are often governed by national rules. However, giving serious thought to the points contained this guidance document and ensuring that all elements are covered in the ToR should help commissioning authorities to select evaluators who will be clear on what precisely they must deliver and the expectations concerning the quality of those deliverables.

ANNEX 2

GUIDANCE ON QUALITY MANAGEMENT OF EXTERNAL EVALUATIONS

Building on the European Commission experience, this document provides guidance to assist national and regional authorities in designing and managing evaluations to support the delivery of good quality. A strategy to manage quality should inform the entire duration of an evaluation from its planning through to the dissemination and use of its findings. It relies on internal and external mechanisms, and can entail the following indicative steps:

1. PREPARATION

a. Assignment of a budget and drafting of the Terms of Reference[8]

The quality of an evaluation hinges largely on the precision of its Terms of Reference (ToR), which set the ground for future work. The ToR usually define the objectives of an evaluation, the role and responsibilities of the evaluator and evaluation client, the duration of the contract and the resources allocated to the project. The European Commission (EC) recommends to pay particular attention to the specification of evaluation questions and to the estimation of data requirements (above all at the level of beneficiaries and supported entities), and to ensure that the budget is in line with what is requested. In general, it is good practice to consult with colleagues most competent on the subject at hand and to discuss evolving drafts within the organization to ensure the evaluation addresses questions that can fruitfully contribute to policy debate.

Those responsible for drafting the ToR should reflect on the amount of work involved in the proposed evaluation. This should include reflection on the data available already, that which the contractor will have to gather, how long it will take to analyse the data and how long it will take to draft a good quality analytical report. Evaluations for DG REGIO and EMPL typically last for a year to 18 months. While national or regional evaluations may be shorter in duration, it is important to allow them sufficient time to deliver a good quality product.

Please consult annex 2 on guidance on ToR for impact evaluations for more details.

b. Selection of the tenderer

The selection process is an essential part of managing quality. The EC recommends appointing a **selection committee** responsible for evaluating the bids against the criteria set out in the ToR. It is important that the committee is allocated enough time and resources to carefully assess the proposals. The quality of the selection, however, is not simply a function of the time spent reading the bids, but also hinges on how well-

For further details on this topic, please consult the "Guidance for the Terms of Reference for Impact Evaluations", available online at: http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm#1.

specified the exclusion, selection, and award criteria were in the ToR. The EC discourages focusing excessively on price as the criterion of award. It is advisable to set out clear award criteria and quality requirements, including a clause regulating the early termination of the contract conditional on the quality of the work provided.

2. PROCESSES: GENERAL MANAGEMENT

It is good practice to appoint an **official responsible for the evaluation and key point of contact** with the consultant. In order to ensure continuity of service, it is advisable to designate a second point of contact within the administration. The ToR should require the winning bidder to set out a similar structure. In general, roles, back-ups and chains of responsibility should be clear to everyone involved in the evaluation.

The EC recommends appointing a **Steering Group (SG)** to coordinate the process of the evaluation. It fulfils both an institutional-representative and a technical-methodological function. In the former role, the SG ensures the interests of all major stakeholders/partners are taken into consideration and that the institutions which might have to act on the recommendations are involved. It is also important to ensure that the steering group performs its tasks without conflict of interest. In the technical-methodological function, the SG safeguards the technical quality of the evaluation from a methodological viewpoint and guarantees independence of the evaluation by relying on scientific experts. In its technical capacity, the steering group may advise on the terms of reference; It supports the evaluation work, for instance by facilitating access to the documentation and data required for evaluation purposes; and regularly assesses the quality of the deliverables.

It is important that the steering group ensures evaluation activities are conducted in a professional and ethical manner. This includes that

- there is impartiality, i.e. absence of bias; in particular, evaluators have not been directly responsible for the policy development, design or overall management of the subject of evaluation;
- there should be consultation of stakeholders at all stages of the evaluation process; and
- evaluators are independent, i.e. free to present their results without compromise or interference although they take account of the steering group's comments on evaluation quality and accuracy.

For particularly important or complex evaluations, the technical-methodological function can be further enhanced through the appointment of a **Scientific Committee**, consisting of methodological and topical experts external to the administration or to the consultant. Dedicated expert meetings are organized, during which the experts comment on the appropriateness of the methodology, and its application, and the quality of analysis, highlighting any inconsistencies and suggesting new lines of analysis.

Should unforeseeable situations arise during the contract, the administration, or contracting authority, may wish to issue an **amendment to the contract**; should the contractor fail to meet the requirements set out in the contract and cannot/will not comply after discussions with the contracting authority, procedures for **termination of the contract** can also be started. In the practice of the EC, a formal letter is first sent to the

contractor, with the history of the discussions held and the points that raise serious concerns. The letter clearly mentions the concerns for breach of agreement and the intent to terminate the contract should the points not be resolved and allows for a reasonable amount of time for the contractor to react/comply (this procedure also forms an integral part of the evaluation contract).

3. PRODUCTS: PROJECT MILESTONES AND DELIVERABLES

a. Kick-off meeting and Inception Report

It is good practice to organize a kick-off meeting during which the contractor is briefed about the ToR, with the help of a 4-5 page note outlining the background of the evaluation, practical information regarding data, deliverables and deadlines, as well as feedback on the points to be developed in the Inception Report. This is also the occasion to discuss any unclear elements in the bid or ToR, and to stress the importance of certain evaluation questions.

The Inception Report is a document detailing the methodology the contractor will use in carrying out the tasks of the ToR. It is usually required about one month after the kick-off meeting and it is a crucial step in ensuring an evaluation of quality is delivered. It is advisable that quality control measures assess the following elements:[9]

- All points in the ToR are addressed, as outlined in the original bid;
- Any weaknesses in approach highlighted in the kick-off meeting are addressed;
- The data gathering approach is reasonable and feasible and will deliver the data to answer the main evaluation questions (in particular as to availability of data at the level of beneficiaries/supported entities);
- The balance between desk research and fieldwork will deliver the information to answer the main evaluation questions;
- Appropriate statistical and other methods are proposed for data analysis, whether this is analysis of data obtained from documentation provided by the national administration or data generated by the consultants through surveys or gaining access to administrative data;
- The fieldwork is described and the research methods are appropriate – i.e., types of interview method – online, telephone or face to face, stakeholder interview, focus groups; draft questionnaires can be examined to ensure all questions are included and asked in appropriate way (balance between open and closed questions, absence of bias, etc.), and templates are reviewed;
- The identification of regions, programmes, or projects for case studies is based on appropriate statistical and other analysis.

[9] All deliverables are examined against the ToR and the quality criteria the Commission has defined in its Guide on Evaluating Socio-economic Development (EVALSED), available online at http://ec.europa.eu/regional_policy/information/evaluations/guidance_en.cfm

- In cases where there is more than one consultant working on the evaluation contract, i.e., the winning bid came from a consortium, the coordination mechanism among the consortium members should be set out clearly, including for example quality control procedures for each deliverable.

The EC advises that the steering group and the external experts, if relevant, review the draft Inception Report, as well as any other draft report delivered in the course of the contract. A formal note outlining the conclusions of the steering group and any additional requirements in order to comply fully with the requirements of the ToR should be sent to the contractor. Acceptance of any report on the part of the contracting authority, as well as any payment (usually linked to the submission and acceptance of important deliverables) is typically conditional on compliance with what is prescribed in the formal note.

b. Intermediate Reports and monthly Progress Reports

The EC has found it to be helpful to include the delivery of intermediate reports during evaluation contracts, particularly those which last longer than 6 months. This means that the contracting authority gains insights to the quality of work underway and can intervene if it is of insufficient quality. It is risky to ask to see nothing between the Inception Report and the draft Final Report when it may be too late to fix an evaluation which has gone in the wrong direction.

Intermediate reports provide an overview of the work carried out by due dates in accordance with the ToR (preliminary conclusions and results of specific tasks). Quality checks should follow a process similar to that for the Inception Report[10]. Particular attention should be paid to the quality of literature reviews and pilot case studies, when applicable. Does the review cover existing literature on the topic and is it analysed appropriately? Have appropriate interview tools been used in case studies, were all the appropriate questions included and were all relevant stakeholders interviewed? All arguments should be evaluated and the presentation of supporting evidence verified. This is an essential quality control, as the quality of the literature review and the pilots will form the basis for the quality of all future deliverables.

Progress Reports: Communication is of crucial importance for the quality of the evaluation. It usually occurs informally on a weekly basis (e.g. email exchange). However, formal monthly progress reports, of up to two pages, can help ensure the work is carried out in a timely fashion (the EC requires progress reports in the ToR).

[10] Ibid

c. Final Report, Executive Summary, communication and use of evaluation findings

Close to the end of the contract, a draft final report is presented, containing results, conclusions and sometimes recommendations on all evaluation questions of the ToR.

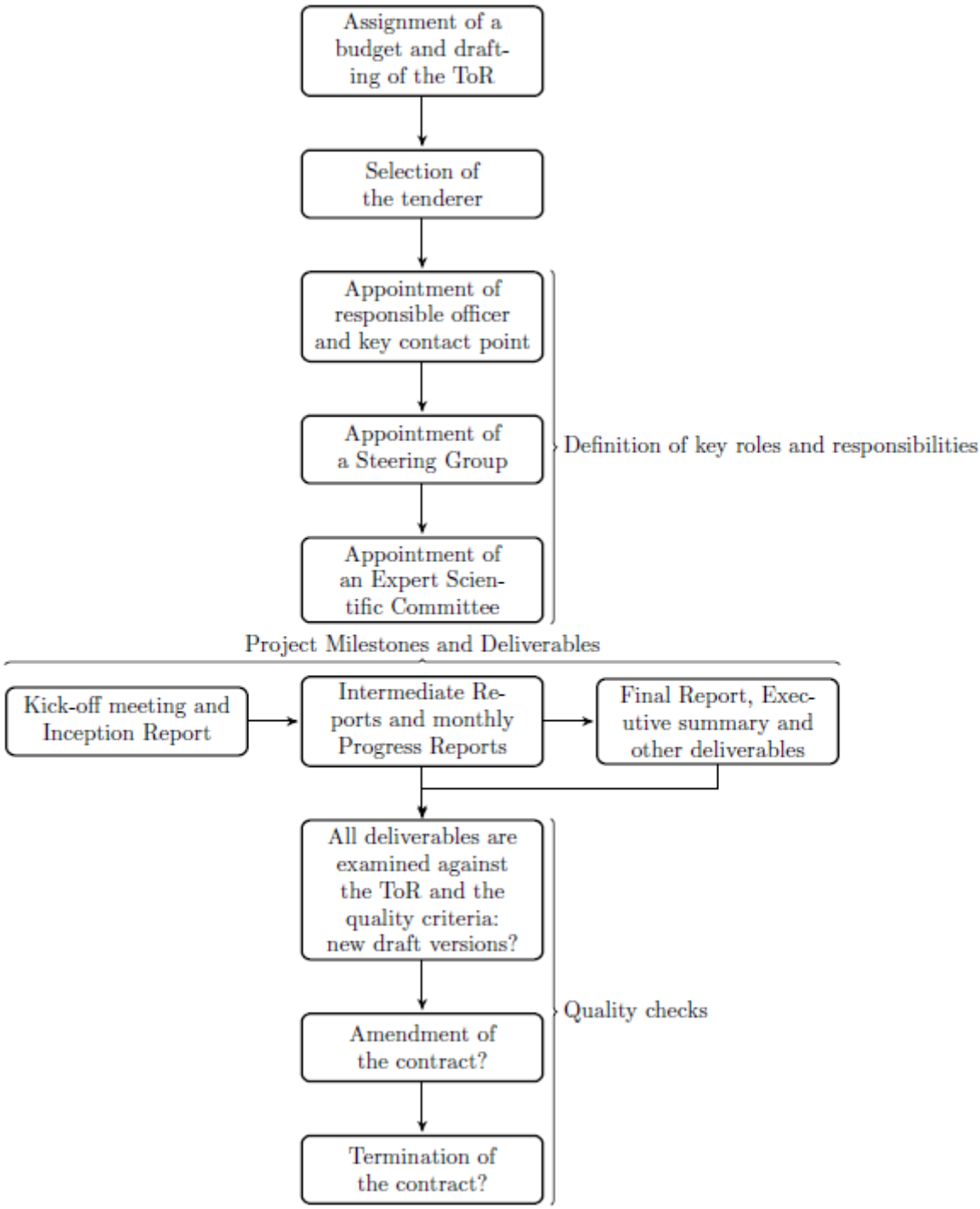
The final evaluation report, as a minimum, should set out the purpose, context, objectives, questions, information sources, methods used, evidence, findings and recommendations. The evaluation methods used must be clearly explained in a methodology section of the report. This should include how the sample was derived, the data collection methods (questionnaires should be included as an annex to the main report), issues encountered in the data collection (biases) and the implication of these on the final results and conclusions. The evaluation report should present in a complete and balanced way the evidence, findings, conclusions and recommendations. It is common practice that the evaluation report includes an executive summary, a synthesis report of the results (particularly relevant in the case of large studies), as well as helping to communicate the use of the results (e.g. preparation of slideshows, organization of workshops and conferences).

As with all deliverables, the EC encourages the organization of a SG meeting to discuss the draft final report and suggest amendments. Amendments and further developments of the report are requested in writing and revised draft final reports are received (as many as necessary until the quality of the reports is satisfactory). All text should be reviewed closely to ensure the logic of arguments and the existence of evidence (quantitative and qualitative) to support conclusions. At this stage, it is advisable that the quality of language is also part of the review.

The use of the results concerns the impact an evaluation can have in an organizational setting, as measured by the changes in ongoing practice, policy, or decision making (i.e. the evaluation as a knowledge resource). This requires a robust understanding of the existing context as regards to variables such as timing of the decision making cycles, organizational structures and processes, or available knowledge management tools. The commissioner of a study therefore needs to develop his/her own conclusions on evaluation findings: not all conclusions might be credible and often recommendations need to be adapted to a changing context. With regard to the latter point, the EC has found it to be good practice to appoint someone in-house responsible for the study and the use of evaluation results. On a broader note, the following table summarizes the main drivers of the effective use of an evaluation.

Driver of use	Description
Characteristic of the receiver/learner	Attitude of the commissioner and his/her capacity to respond to evaluation results and to disseminate them
Type and saliency of the policy	Role and relevance of the evaluation for the commissioner within the policy process
Timing of the evaluation	Positioning of the evaluation in the policy and decision making cycle
Approach and methods chosen for the evaluation	Openness of the evaluation process and involvement of relevant stakeholders
Quality	Informative content of evaluation results and appropriateness of the form of the deliverables

In many cases a single evaluation will not be enough to draw firm policy conclusions. Building up knowledge from several evaluations and using other sources is essential. For the purpose of organizational learning, it is particularly important to ensure knowledge of results is delivered to those in the administrative or political body mostly concerned in their activity by the findings of the evaluation. More specifically, it is advisable to ensure that the final report and the main conclusions of the study are made available to the relevant monitoring committee. It is also good practice to publish evaluations. Ensuring availability of evaluations to the public is a requirement in the 2014-2020 period.


Assignment of a budget and drafting of the ToR
Selection of the tenderer
Appointment of responsible officer and key contact point
Appointment of a Steering Group
Appointment of an Expert Scientific Committee
Definition of key roles and responsibilities
Project Milestones and Deliverables
Kick-off meeting and Inception Report
Intermediate Reports and monthly Progress Reports
Final Report, Executive summary and other deliverables
All deliverables are examined against the ToR and the quality criteria: new draft versions?
Amendment of the contract?
Termination of the contract?
Quality checks