



EUROPEAN COMMISSION
Directorate General Regional and Urban Policy
Policy
Evaluation and European Semester

The Programming Period 2014-2020

GUIDANCE DOCUMENT ON MONITORING AND EVALUATION
– EUROPEAN REGIONAL DEVELOPMENT FUND AND COHESION FUND –

Guidance for the Terms of Reference for Impact Evaluations

October 2013

Quotation is authorised as long as the source is acknowledged.

1. 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 draft Regulation (EU) XX/2013 on common provisions on the ERDF, the ESF, the Cohesion Fund, the EAFRD and the EMFF (art. 47 - 49) 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.

2. 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.

3. 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.

- 3.1 Background and context
- 3.2 Evaluation purpose and target audience
- 3.3 Evaluation objective and scope
- 3.4 Evaluation questions and tasks
- 3.5 Approach and Methodology
- 3.6 Timing and deliverables
- 3.7 Evaluation team composition and required competencies
- 3.8 Management arrangements
- 3.9 Budget and Payment
- 3.10 Proposal Submission

3.1. 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.

3.2. 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.

3.3. 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.

3.4. 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 questions¹:

- What change can be observed in relation to the objectives of the intervention?
- To what extent can observed changes be attributed to the intervention?

¹ 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.

- Are there unintended impacts?
- What mechanisms delivered the impact? What are key contextual features for these mechanisms?

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 need 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.

3.5. 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., discontinuity design, propensity score matching, etc.).

3.6. 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).

3.7. 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.

3.8. 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.

3.9. 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).

3.10. 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.

4. 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.