



European Commission



EU-ESPA





Terms of Reference for evaluations

David Alba
DG REGIO Evaluation
Unit

ToR?

What is the Terms of Reference

- Aspects of how an evaluation will be conducted
- Presents the objectives of the evaluation
- Role and responsibilities
- Resources available

Double purpose

- Hire the evaluator on a competitive basis (public procurement)
- To have high quality evaluations (this depends on how accurate and well-specified the ToR is.)
- It is a contractual arrangement.
- Inherent problem with double role
 - should enable selecting a competent evaluator
 - should also ensure that the contractor delivers

What is the Terms of Reference

- Aspects of how an evaluation will be conducted
- Presents the objectives of the evaluation
- Role and responsibilities
- Resources available

Double purpose

- Hire the evaluator on a competitive basis (public procurement)
- To have high quality evaluations (this depends on how accurate and well-specified the ToR is.)
- It is a contractual arrangement.
- Inherent problem with double role
 - should enable selecting a competent evaluator
 - should also ensure that the contractor delivers

Typical Structure



Typical Structure 1

- Evaluation purpose and target audience
- Evaluation objective and scope
- Evaluation questions and tasks
- Approach and Methodology
- Timing and deliverables

Typical Structure 2

- Evaluation team composition and required competencies
- Management arrangements
 - Budget and Payment
 - Proposal Submission

Typical Structure 1

- Evaluation purpose and target audience
- Evaluation objective and scope
- Evaluation questions and tasks
- Approach and Methodology
- Timing and deliverables

Typical Structure 2

- Evaluation team composition and required competencies
 - Management arrangements
 - Budget and Payment
 - Proposal Submission

Most relevant sections

Background and context

- 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

Evaluation objective and scope

- Objective of the evaluation reflects what the evaluation aims to find out:
- no more than two or three objectives
 - few issues, deeply rather than examine a broader set superficially
- Scope defines the focus of the evaluation
- time, period, the geographical and thematic coverage of the evaluation, the target groups and the issues to be considered;
 - evaluation must be realistic given the time
 - and resources available

Evaluation tasks and questions

- Work it wants the evaluator to undertake should be structured into:
- A series of evaluation tasks with a logical sequence to build up evidence on the subject being evaluated;
 - Under each evaluation task there will be specific evaluation questions;
 - It should represent a real need for knowledge, understanding or identification of new solution

Evaluation question considerations

- 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.

Main impact evaluation questions

- 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 factors for these mechanisms?
- Does the impact vary by sub-group within the main target group?
- Will short-run effects of the intervention differ from those in the long run?

Background and context

- 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

Evaluation objective and scope

Objective of the evaluation reflects what the evaluation aims to find out:

- no more than two or three objectives
- few issues deeply rather than examine a broader set superficially

Scope delimits the focus of the evaluation

- time, period, the geographical and thematic coverage of the evaluation, the target groups and the issues to be considered.
- evaluation must be realistic given the time
- and resources available

Evaluation question considerations

- *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.*

Evaluation tasks and questions

Work it wants the evaluator to undertake should be structured into:

- A series of evaluation tasks with a logical sequence to build up evidence on the subject being evaluated.
- Under each evaluation task there will be specific evaluation questions.
- It should represent a real need for knowledge, understanding or identification of new solution

Main impact evaluation questions

- 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?

Approach and Methodology

Two possibilities:

You 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,
- 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).

Main points

- You need to have an understanding of the potential, limits and basic technical features of methods
- You 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
- To prevent this, the ToR could request further detail to be added in the tender documentation (e.g., description of how the methods will be assessed etc.).

Counterfactual considerations

- A ToR should include information on:
- data available to construct a comparison or control group
 - assessment of the quality of this data
 - analyse strengths and weaknesses and propose which among various techniques will be used (e.g., which discontinuity design(s), which forms of matching, etc.).

Impact evaluation considerations

- Start with a review of the theory of change which underpins the intervention being evaluated
- ToR should ask for an outline of what data sources will be used for this

Theory based approach considerations

- Theory-based evaluations should answer questions concerning how and why a programme has, or has not, had an impact
- investigate the causal linkage that relates inputs, activities and outputs to impacts
- scrutinise all the assumptions that underlie this 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.

Two possibilities:

You 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,
- 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).

Main points

- You need to have an understanding of the potential, limits and basic technical features of methods
- 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.
- To prevent this the ToR could request further detail to be added in the tender documentation (e.g., description of how the methods will be combined etc.).

Impact evaluation considerations

- Start with a review of the theory of change which underpins the intervention being evaluated
- ToR should ask for an outline of what data sources will be used for this

Theory based approach considerations

- Theory-based evaluations should answer questions concerning how and why a programme has, or has not, had an impact
- investigate the causal linkage that relates inputs, activities and outputs to impacts.
- scrutiny all the assumptions that underlie this 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.

Counterfactual considerations

A ToR should include information on:

- data available to construct a comparison or control group.
- assessment of the quality of this data
- analyse strengths and weaknesses and propose which among various techniques will be used (e.g., which discontinuity design(s), which forms of matching, etc.).



Further considerations

Do we always need CFI evaluations?

- Yes, Ideally counterfactual and theory based approaches should complement each other.
- However you cannot always perform counterfactual evaluations? See ETC interventions!
- With proper narrative naive before-and-after comparisons may be used.
- Theory based evaluations also look at data and address the quantification of effects.
- Policymakers should use the results of both sets of methods as they see fit.

Is there an ideal evaluation guaranteeing all valid answers?

- All methods and approaches have their strengths and weaknesses.
- They should be adapted to the specific question to be answered, to the subject of the programme and its context.
- Whenever possible, evaluation questions should be looked at from different viewpoints and by different methods. This is the principle of triangulation.
- The costs of evaluation need to be justified by the possible knowledge gain.
- When deciding on an evaluation what is already known about an intervention needs to be considered.
- There is no "best" method for all circumstances.

Do we always need CFI evaluations?

- Yes, Ideally counterfactual and theory based approaches should complement each other.
- However you cannot always perform counterfactual evaluations? See ETC interventions!
- With proper narrative naive before-and-after comparisons may be used.
- Theory based evaluations also look at data and address the quantification of effects.
- Policymakers should use the results of both sets of methods as they see fit.

Is there an ideal evaluation guaranteeing all valid answers?

- All methods and approaches have their strengths and weaknesses.
- They should be adapted to the specific question to be answered, to the subject of the programme and its context.
- Whenever possible, evaluation questions should be looked at from different viewpoints and by different methods. This is the principle of triangulation.
- The costs of evaluation need to be justified by the possible knowledge gain.
- When deciding on an evaluation what is already known about an intervention needs to be considered.
- There is no "best" method for all circumstances.

***Don't pay for something you
already have!***

Ex-ante evaluations

- The ex ante evaluation of programmes can be understood also as a theory-based analysis (if properly implemented), assessing the strength of the theory of change and the logical framework before the programme is implemented.
- The task is to look into different interventions and to confirm the functioning of the logic.

Any other evidence?

- Evaluations carried out during the preceding periods
- EU ex post evaluations
- Evaluation literature and preparatory studies
- ESPON/URBACT/Interact studies

Ex-ante evaluations

- The ex ante evaluation of programmes can be understood also as a theory-based analysis (if properly implemented), assessing the strength of the theory of change and the logical framework before the programme is implemented.
- The task is to look into different interventions and to confirm the functioning of the logic.

Any other evidence?

- Evaluations carried out during the preceding periods
- EU ex post evaluations
- Evaluation literature and preparatory studies
- ESPON/URBACT/Interact studies

Key success factors for effective impact evaluations

Timing

- As late/as early as possible: depends on nature of intervention; consider evaluating similar interventions 2007-2013
- All impact evaluations will not be carried out at the same time

Data planning and availability

- Supported entities and individual participants
- Data from other sources of information than monitoring systems: unemployment records, tax records...
- Counterfactual: data for non-treated groups (probably not feasible for all ETC programmes)
- Need to be planned in advance

Contractors and budget

- Expectations should be properly managed! Beware when defining evaluation questions.
- Evaluation are very dependent on the quality of the evaluator.
- Available budget always limits the scope!
- Somebody understanding basic evaluation terms is needed within the organisation.

Timing

- As late/as early as possible: depends on nature of intervention; consider evaluating similar interventions 2007-2013
- All impact evaluations will not be carried out at the same time

Data planning and availability

- Supported entities and individual participants
- Data from other sources of information than monitoring systems: unemployment records, tax records...
- Counterfactual: data for non-treated groups (probably not feasible for all ETC programmes)
- Need to be planned in advance

Contractors and budget

- Expectations should be properly managed!
Beware when defining evaluation questions.
- Evaluation are very dependent on the quality of the evaluator.
- Available budget always limits the scope!
- Somebody understanding basic evaluation terms is needed within the organisation.

Questions 1

- How to define quality assessment criteria, so they are objective and can stand audit control and in the same time tell us something about suppliers' quality and competences?
- Practical tips how to make ToR for Programme evaluation
- What qualification criteria should be used? What award criteria are used in practice? Is experience of evaluators taken into consideration? How to ensure quality of reports through ToRs? How to approach operational evaluations to be useful for the impact evaluations?

Questions 2

- Under which conditions or circumstances is it most relevant to have all evaluations under one ToR and/or in several ToR?
- Have there been programmes that have had both operational and impact evaluations within the same ToR? If yes - how has it worked so far? Do programmes have multiple parts to their operational evaluations and if yes - have these gone in the same ToR?
- When ToR are considered successful? What is the absolutely necessary to take into consideration?

