



European
Commission

How-to Guide: **Evaluation Questions**

The evaluation questions are the backbone of an evaluation. This guide will help you understand why, and how to develop critical questions

Prepared by the evaluation services
of DG INTPA, DG ENEST, DG MENA and FPI
Brussels, June 2025



Purpose of this How-to Guide: This guide is designed to assist evaluation managers in understanding the significance of crafting effective indicative evaluation questions to be included in their Terms of Reference. It outlines the key considerations for developing various types of questions and provides steps for their drafting.

Target audience: This guide is specifically prepared for evaluation managers preparing their evaluation Terms of Reference (ToR) under the FWC SEA 2023 – Lot 16. However, other evaluation stakeholders, including the M&E focal points and partners could also benefit from the guidance provided.

How to use this How-to Guide: We suggest that you navigate through the key points provided in this guide alongside the Terms of reference you are working on ([see also ToR template and guidance for intervention level evaluations - SEA FWC SharePoint page, Section 2.2, Lot 16](#)).

INTRODUCTION

Evaluation questions (EQs) set a strong foundation for an evaluation. EQs are important as they shape the evaluation's direction, influence the methodology that the evaluators will develop, and determine the findings of the evaluation.

Every evaluation is unique, therefore there is no predefined set of EQs that is valid for all. The choice of questions will depend on multiple factors including the type of evaluation, its scope and purpose, knowledge needs and prior evaluative knowledge.

The drafting of the indicative EQs to be included in the ToR is an iterative process.

1 > Consider who will use the evaluation, what will they use it for, what they need to know, and the resources required. Ask yourself these four questions

- I. **Why is the evaluation needed?** For example, is it for accountability, learning, or for both.
- II. **What do you want to know?** Look back at the intervention logic (or logframe) of the intervention(s) to be evaluated. Are there any specific aspects, components or results (or links between expected results) that you would like the evaluation to focus on? For example, *Were the intended outcomes achieved? What factors played a greater role in the achievement of a particular outcome?*
- III. **What would you like to learn from this exercise?** Look at the target groups, stakeholders, and other key actors. You might want to explore the issues of ownership, participation, governance of the action. For example, *did the intervention strengthen institutional and/or human capacities of the target groups (were there differences among groups)? And what are the chances that the strengthening of these capacities will lead to improved performance in the short medium to long term? Did the governance mechanisms of the intervention work well?*
- IV. **How are you going to use the evaluation?** For example, is it to adjust implementation, or to inform future planning? Who would be the key users of your evaluation within and outside the EU institutions?



Useful Tips: Write down the answers to these questions in the form of list of issues to be addressed by the evaluation, using simple words. Keep an open mind, come up with a long list of issues (be curious, be specific), you can then reduce the list and fine-tune it later.

2 > Engage your Reference Group members¹ and other stakeholders early in the process

Share your considerations with the Reference Group and discuss with them. Verify if there are other issues they would like to be assessed by the evaluation. This would be helpful in the development of relevant, credible, and useful questions. Identifying evaluation questions is a shared responsibility. By gathering inputs from a wide range of stakeholders, evaluations can better reflect diverse perspectives and experiences. This inclusive approach not only enhances the acceptance and implementation of evaluation recommendations but also increases the likelihood of findings being used for learning, decision-making, and action.



Useful Tips: Make sure that the Reference Group and other stakeholders review and own the questions before finalising the ToR, ensuring they anticipate the answers.

¹. See also [How to Guide on Evaluation Reference Group](#).

3 > Prioritise your EQs and limit the EQs to a maximum of 10. Often, 5-6 well-targeted EQs are sufficient!²

Too many evaluation questions will reduce the time evaluators have to properly address each one of them, negatively impacting the usefulness of your evaluation report. As a practical consideration, the wider the focus of the analysis, the higher the cost of the evaluation will be.

What if you have more than 10 key issues you would like the evaluation to focus on? In this case, consider grouping or clustering related issues (for example, examining the same main issue from different dimensions). If you still have more than 10 key issues, consider prioritising them and select the ones that are most important to you and your Reference Group.



Useful Tips:

- > Remember, the evaluation team will further develop the judgement criteria³ to answer the EQs thereby there is a scope for addressing specific aspects and further defining the scope.
- > You may want to ask evaluators to further refine the EQs in their Specific Contract O&M. This is a good way to check if they have understood the objective of the requested evaluation.

4 > Organise your EQs according to different criteria

> By **evaluation criteria** (6 DAC + the EU Added Value). In this case, each selected evaluation criterion should be covered by at least one Evaluation Question; preferably without mentioning it in the evaluation question.⁴

> By **clusters** (this is the option recommended in the case of DG ENEST and DG MENA) covering:

- **Transversal areas** such as i) policy framework and responsiveness, ii) management and governance (institutional set-up), iii) EU cooperation potential (Team Europe approach) and EU added value, iv) partnerships (engagement, co-ordination, and complementarity with key stakeholders at local, regional, national and/or international level). Define the transversal areas that are most relevant to your evaluation if you choose this option.
- **Thematic areas**. In this case, the EQ would typically focus on the chain of results leading to achievements in a specific **sector** or sub-sector in line with intervention logic of the intervention(s) to be evaluated. This type of question allows to assess the links between different levels of results eliciting greater cohesiveness of the answer.



Useful Tips:

- > If EQs are organised by transversal and/or thematic areas, one or more evaluation criteria would be covered at the same time within each area. This choice often allows to bring into greater focus what it is you want to know as you are not tied to pre-defined categories, issues and/or evaluation criteria.
- > Regardless of how you decide to organise your EQs, ensure that all selected evaluation criteria are covered. Remember that it is not necessary to cover all the DAC criteria, but the EU Added Value criterion should not be omitted. For example, if your mid-term evaluation has already assessed Relevance, you may exclude this criterion in the final evaluation to focus more on the other evaluation criteria.
- > Make sure you number your EQs; this will simplify your interaction with the evaluators during their finalisation and reporting.

2. In the case of DG ENEST and DG MENA services, the number of EQs is expected to be limited to 5-6 so as to really target key issues at stake where thorough analyses and answers are expected. The aim is to avoid having evaluation studies that are encompassing many issues and that result in broad/superficial insights/answers, and thus in useless (and costly) evaluations.

3. The **judgement criteria** (JCs) each specify a different aspect of the evaluation questions and look into the merits or success of the evaluand (single intervention, group of interventions under a specific area). They are formulated as positive statements and are used to answer an evaluation question positively or negatively. To some extent, they can be defined as well as sub-questions of the main headline. There usually should be more than one JC per evaluation question, usually 2-3 but can also be 4, in any case their number (per evaluation question and overall) should be manageable.

4. The Better Regulation Guidelines sets as mandatory the following five criteria that have to be addressed in each evaluation: Relevance, Efficiency, Effectiveness, Coherence and EU added value. If one or more criteria are not covered, due justification must be provided. See also section 2.2 and annexes on the Evaluation Criteria in the [ToR template and guidance for evaluations of interventions under SEA 2023 \(Lot 16\)](#). For further details see also the OECD DAC guidance document, which can be accessed online [HERE](#).

5 > Design focused and open-ended EQs

Ensure the questions are **clear, directly relevant to the evaluation objectives**, and will **provide actionable insights**. Avoid generic formulations, and tailor your EQs to the issues at stake in the evaluand. For example, the EQs can come from the key evidence you are seeking emerging from the logframe, instead of generic references to outputs, outcomes, or impact.

Prefer open-ended to closed-ended questions to have robust inquiry, and as a result detailed responses by the evaluators in the evaluation report.

Transforming an issue into an EQ involves several steps:

For example, as an evaluation manager, during the initial reflection or brainstorming with your reference group, you might have identified a broad issue such as: *“Did the intervention deliver/achieve its expected results?”* This is a closed question, so it should be rephrased into an open question to facilitate a more detailed exploration, for example: *“To what extent have the expected results been delivered/achieved?”* However, the term “result” is too generic as it encompasses outputs, outcomes, and impacts. Therefore, you need to specify the results you are referring to: *“To what extent has the delivery of the outputs contributed to the achievement of the intended outcomes?”*

Even this question remains somewhat generic and could apply to many different evaluations. To make it more specific to the intervention, context, or evaluation, you need to tailor the formulation. Depending on the intervention or sector, it could become more focused, such as:

- i) *“To what extent have EU interventions supported legal reforms promoting closer alignment to international standards, norms, and values in Rule of Law?”* or
- ii) *“To what extent has EU support helped the partner country’s Micro, Small, and Medium Enterprises (SMEs) to position value-added goods on national and international markets?”*



Useful Tips:

- > Ensure consistency among the evaluation objectives, its scope, and the Evaluation Questions.
- > Use straightforward, and plain language without jargons.

6 > Select different Types of EQs

As an evaluation manager, it is essential to understand the various types of EQs. Below you can find three different types of EQs that you can use in your ToRs:

> **Descriptive Questions** ask what has happened and require evaluators to define, observe, and measure change, often from the point of view of various stakeholders. They address changes, be they expected or unexpected, directly, or indirectly linked to the intervention. Lines of inquiries are:

- What was the situation before the intervention (baseline) and what is the situation now?
- How do changes differ for each area/sector/affected group?

These questions help evaluate relevance, efficiency, effectiveness, and EU added value by providing detailed descriptions using multiple data sources.

> **Causal Questions** explore whether and how an intervention works, for whom, and under which circumstances. They include testing clear hypotheses, such as:

- Has the intervention produced xx changes (or stopped change?) and how? What contributed to the changes?
- How and to what extent the provision of electricity in the community X impacts on gender equality?

These questions help evaluate Impact and Effectiveness, by using mainly experimental or quasi-experimental designs to produce knowledge that can be used to improve interventions.

> **Normative Questions** assess how an intervention fares against specific criteria. EMs should state the rationale for their chosen criteria:

- Is the intervention worth the cost?

Answering these require systematic, transparent, and defensible approach to avoid arbitrary judgments. These questions help answering Relevance, Coherence, and EU added value criteria.

7 > Examples of evaluation questions

Here are some examples of EQs presented by key theme / area of focus to illustrate how different types of evaluation questions might look for your evaluation. EQs can be organised according to different criteria: i) by **evaluation criteria**; ii) by clusters covering **transversal areas**; iii) by **thematic areas**. The three structuring criteria are not mutually exclusive, you can have a mix of evaluation questions (thematic, transversal or focused on one evaluation criterion) but ensure that all selected evaluation criteria are covered by the final set of indicative questions drafted.

Use the examples as a guide rather than copying them directly into your ToR. Each evaluation is unique, so tailor your questions by referring to specific results, sectors, target groups, to fit your specific context. Since EQs are the backbone of your evaluation, invest time in designing them with your reference group and other stakeholders, using the guidance and examples provided in this guide.

Please note that we have included gender-responsive questions as standalone items in the structuring criteria in the table below. However, in your specific evaluation, you can either keep them as standalone questions or work with the evaluation team to incorporate them into the Judgement Criteria. This ensures that gender aspects are thoroughly examined and well-presented in the report.

Examples of evaluation questions by area of focus	Structuring criteria	Evaluation criteria covered
Responsiveness to policy priorities and needs		
These EQs examine the design and relevance of interventions over time. The different examples, cover to a smaller or larger extent the intervention's alignment with government and EU goals, adaptivity and responsiveness to target groups' needs, and inclusion of cross-cutting themes and marginalised groups.		
<i>To what extent was the design of the intervention relevant to the national / sector policy priorities and to the EU priorities in the sector considering the changing circumstances during the implementation period?</i>	Evaluation criteria	Relevance
<i>To what extent has the intervention design considered other on-going interventions in the same sector and incorporated lessons learned from the experiences of the EC and other donors, concerning factors that facilitate or impede the success of support in the partner country?</i>	Transversal	Relevance, Coherence
<i>To what extent does the intervention (still) meet the needs of its target groups and support relevant government policies?</i>	Evaluation criteria	Relevance
<i>To what extent was gender analysis considered during the formulation of the intervention, including the intersection of gender with other forms of inequality (e.g., ethnicity, age, sexual orientation)?</i>	Evaluation criteria & Gender-responsive	Relevance
<i>To what extent was gender equality considered and included throughout the intervention (design, implementation and monitoring) and how was it consistent with national policies or international instruments on gender equality and relevant international human rights obligations?</i>	Transversal & Gender-responsive	Relevance, Efficiency, Coherence
Coherence, coordination and EU added value		
These EQs examine the compatibility of the intervention with other interventions in a country, sector, or institution, to:		
i) determine whether the intervention complements and supports other efforts, avoiding duplication and enhancing overall impact; and / or		
ii) assess the extent to which the intervention brings additional benefits beyond what would have resulted from Member States' interventions alone in the partner country.		
<i>To what extent was the intervention designed and implemented to complement other EU (and non-EU) interventions?</i>	Evaluation criteria	Coherence
<i>To what extent have the results of the intervention complemented/been supported by other EU interventions in the area of external action and foreign policy?</i>	Transversal	Coherence, Effectiveness
<i>What is the additional value resulting from the EU intervention), compared to what could be achieved by others (i.e. EU Member States, other donors, and the country/region) and what are the reasons for this?</i>	Evaluation criteria	EU Added Value
<i>To what extent – and how – does the intervention add benefits to or link to Member States' interventions in the area of gender mainstreaming/gender equality?</i>	Thematic (gender)	EU Added Value
<i>To what extent has the EU intervention been designed and implemented so as to maximise the European (i.e. Commission + EEAS + EU Member States + EU agencies) cooperation potential and the EU added value?</i>	Transversal	EU Added Value, Coherence
<i>To what extent was the intervention coherent with EU commitments and strategies in the area of gender mainstreaming/gender equality thereby contributing to the implementation of the EU Gender Action Plan?</i>	Transversal & Gender-responsive	Coherence, Efficiency, Effectiveness
<i>To what extent has the coordination and sharing of roles between the EC and Member States contributed to optimise the impact of the support?</i>	Transversal	Coherence, Effectiveness, & Impact

Examples of evaluation questions by area of focus	Structuring criteria	Evaluation criteria covered
Efficiency: governance mechanisms, responsiveness of choice of implementation mechanisms to needs and absorption capacities of target groups		
This set of EQs examines the extent which the intervention delivers, or is likely to deliver, results in an economic and timely manner. The different examples also cover to a smaller or larger extent whether the results delivered respond to the needs of the target groups thereby laying grounds for sustainability, as well as the governance and monitoring mechanisms put in place to ensure smooth implementation of the intervention.		
<i>To what extent - and - how have the Commission's intervention(s) made it possible to produce the expected effects with the lowest transaction costs possible?</i>	Evaluation criteria	Efficiency
<i>To what extent - and how - has the intervention contributed to build/strengthen technical, organisational and financial capacities of its target groups (local CSOs // local governments // xxx)?</i>	Evaluation criteria	Efficiency
<i>To what extent - and how - has the intervention contributed to build/strengthen technical, organisational and financial capacities of its target groups (local CSOs // local governments // xxx) in light of existing needs and absorption capacities?</i>	Transversal	Efficiency, Relevance
<i>To what extent were the resources (financial, time, people, technical and gender expertise) sufficient to address the gender inequalities defined during the formulation of the intervention and how were they spent or allocated to target the structural causes of inequality?</i>	Thematic & Gender-responsive	Efficiency
<i>To what extent have the coordination & internal monitoring systems/ processes put in place at different levels ensured the tracking of progress (including via the integration and use of gender analysis) and facilitated the adoption of measures to ensure smooth implementation?</i>	Transversal & Gender-responsive	Efficiency, Relevance
<i>To what extent has the intervention mainstreamed gender in its activities and resulted in more gender-responsive and gender equal civil society activities?</i>	Transversal & Gender-responsive	Efficiency, Effectiveness
Results and sustainability of results of the intervention(s)		
These EQs examine the extent to which the intervention has achieved, or is expected to achieve, its results; whether it has generated, or is expected to generate, significant positive or negative, intended or unintended, higher-level effects; and whether the net benefits of the intervention are likely to continue over time. Some of the examples provided also assess long-term viability, considering factors such as institutional support, community ownership, and ongoing funding.		
<i>To what extent have EU interventions supported legal reforms promoting closer alignment to international standards, norms, and values in Rule of Law?</i>	Thematic	Efficiency, Effectiveness
<i>To what extent and how has the intervention contributed to improve the (legal, policy, financial) conditions (the so-called 'enabling environment') under which CSOs // SMEs // XXX operate and work?</i>	Thematic	Efficiency, Effectiveness, Sustainability
<i>To what extent did the intervention contribute to achieving its expected results (to be substituted by the respective outcomes/specific objectives), respectively for men/boys and women/girls, and for those marginalised or in a vulnerable situation, also considering an intersectional perspective, where relevant?</i>	Evaluation criteria & Gender-responsive	Effectiveness
<i>To what extent has the intervention contributed to enhance youth empowerment, engagement and participation in public life?</i>	Evaluation criteria	Effectiveness
<i>To what extent and how has the intervention contributed towards durably remedying the backlog in road network maintenance?</i>	Thematic	Efficiency, Effectiveness, Sustainability
<i>Was an appropriate exit strategy planned and implemented and to what extent and how were local partners and various beneficiaries involved in defining and implementing the exit strategy?</i>	Evaluation criteria	Sustainability
<i>To what extent have the EU interventions contributed to improvements in sustainable production practices and with what effects in terms of overall production (variety and quantity)?</i>	Thematic	Effectiveness, Impact, Sustainability
<i>To what extent has the intervention promoted sustainable changes in the area of gender mainstreaming/gender equality and changes in gender power relations?</i>	Transversal & Gender-responsive	Effectiveness, Impact and Sustainability
<i>To what extent do the partners of the intervention take ownership of its results, and how committed are they to sustaining these results after the intervention ends?</i>	Transversal	Relevance, Effectiveness and Sustainability
<i>To what extent has the governance, management and disbursement system of the sub-grant component contributed to the strengthening of both short and long-term project implementation capacity of the grant beneficiaries?</i>	Transversal	Efficiency, Effectiveness, Sustainability
<i>To what extent is EU support contributing to improving private-sector productivity and competitiveness and helping the partner country's Micro, Small and Medium Enterprises (MSME) to position value-added goods on national and international markets effectively?</i>	Thematic	Effectiveness, Impact, Sustainability

If you would like to receive specific support or advice on your evaluation, please contact the Evaluation Help Desk at:

DG INTPA: INTPA-EVALUATION-SUPPORT@EC.EUROPA.EU

DG ENEST: ENEST-PERFORMANCE@EC.EUROPA.EU cc: EVALUATION-SUPPORT@MELDEA.EU

DG MENA: MENA-EVAL-MONITORING@EC.EUROPA.EU cc: EVALUATION-SUPPORT@MELDEA.EU

FPI: FPI-EVALUATION@EC.EUROPA.EU cc: EVALUATION-SUPPORT@MELDEA.EU