



United Nations  
Educational, Scientific and  
Cultural Organization

Organisation  
des Nations Unies  
pour l'éducation,  
la science et la culture

Organización  
de las Naciones Unidas  
para la Educación,  
la Ciencia y la Cultura

Организация  
Объединенных Наций по  
вопросам образования,  
науки и культуры

منظمة الأمم المتحدة  
للتربية والعلم والثقافة

联合国教育、  
科学及文化组织

**Internal Oversight Service  
Evaluation Section**

**IOS/EVS/PI/127  
Original: English**

# **Guidelines for Developing Terms of Reference for External Evaluations**

*April 2013*

## Introduction

These guidelines aim to assist the different entities in the UNESCO system (mainly Sectors, Field Offices and Category I Institutes) in the preparation of Terms of Reference (ToR) for external evaluations<sup>1</sup> of Regular Programme and Extrabudgetary activities. In addition, these guidelines can also be used as a basis for developing the ToR for an external validation exercise of final narrative reports.<sup>2</sup>

An evaluation cannot be effectively carried out without a sound ToR that clearly sets out the framework of an evaluation, the expectations of different stakeholders, and what is to be delivered by whom, when and how. The ToR are a “statement of the background, objectives, and purpose of an evaluation; the individual duties and responsibilities of members of an evaluation team; and timelines for deliverables” (Morra and Rist, 2009). The concerned Sector<sup>3</sup> is responsible for drafting the ToR.

The process for developing the ToR should be a participatory one that involves the key stakeholders and users of the evaluation. This can be guaranteed for example by establishing an evaluation reference group,<sup>4</sup> which provides advice and quality assurance throughout the entire evaluation process. The Internal Oversight Service (IOS) can provide additional advice and feedback on the draft ToR. For evaluations of high strategic importance to the Organization or covering a programme/portfolio/framework representing substantial financial resources, it is recommended to involve IOS in the process of overseeing the quality of evaluations.

---

<sup>1</sup> In the document *Guidance Note on the Evaluation of UNESCO's Extrabudgetary Activities* external evaluation is defined as: “an evaluation conducted by independent entities [including UNESCO's Internal Oversight Service] and/or individuals outside the donor and UNESCO elected through a competitive process resulting in an external evaluation report”.

<sup>2</sup> See the *Guidance Note on the Evaluation of UNESCO's Extrabudgetary Activities*. Please note that external validation exercises are less elaborate and costly than external evaluations. Consequently, the ToR for the former do not need to be as elaborate as presented in this document.

<sup>3</sup> For simplicity purposes, in this document the term ‘Sector’ also refers to Field Offices and Category I Institutes.

<sup>4</sup> This group is composed of representatives of stakeholders (e.g. Sector, external donors, implementing partners, IOS). The purpose of the reference group is to provide guidance and feedback to the evaluation team (and the UNESCO staff managing the evaluation) on (e.g.) the ToR, the selection of the evaluation team, the planning and implementation of the evaluation, and the (draft) report.

# Template for the Terms of Reference

The following structure for ToR of external evaluations is recommended:

- Title
- Background
- Purpose
- Scope
- Methodology
- Roles and Responsibilities
- Qualifications Evaluation Team
- Deliverables and Schedule
- References

## TITLE

The title should include the full name of the evaluand and possibly the type of evaluation.

## BACKGROUND

This section should include a brief description of:

- The history and current status of the evaluand (e.g. the programme),<sup>5</sup> including its objectives, main activities, budgetary information, duration.
- The alignment of the programme with UNESCO's mandate (e.g. the C/5, C/4, Strategic Objectives, Global Priorities); the alignment with other relevant global, regional or national strategies.
- The main stakeholders of the programme; the partnerships and frameworks of collaboration; the relationship between the programme and past/future interventions.
- The rationale for the evaluation (i.e. what prompted the evaluation; who initiated the evaluation and why).

## PURPOSE

This section should include the following elements:

- The purpose of the evaluation: what is/are the main objective(s) of the evaluation (e.g. to assess the effectiveness and relevance of programme x, to develop a better understanding of how and to what extent programme x is facilitating target groups y in achieving z).
- The expected uses of the evaluation by key stakeholders, e.g.:
  - o For accountability purposes (e.g. generating evidence of key achievements and challenges to inform decision-making processes of donors/ UNESCO's Senior Management/Executive Board).
  - o For learning purposes (e.g. generating an understanding of which delivery mechanisms work best under which circumstances and how they can be improved in the future).

## SCOPE

Key elements to be covered in this section are:

- The key dimensions, evaluation criteria and/or questions to be covered by the evaluation.

---

<sup>5</sup> The evaluand can be an activity, project, programme, portfolio, strategy, policy, theme or institutional entity. In this document we will use the term programme.

- Issues of delimitation and coverage (e.g. the timespan of the programme to be covered, geographical coverage, thematic coverage, delimitation issues of the programme).

Regarding the first issue, evaluations can broadly be structured around the five evaluation criteria as defined by the OECD-DAC: relevance, efficiency, effectiveness, impact, sustainability. Indicative questions for each of these criteria are presented in the Box below. It is important to emphasize that an evaluation does not necessarily need to cover all five criteria. Depending on the resources available for evaluation, the principal objective(s) of the evaluation and the use by stakeholders, clear choices on the need for inclusion and relative weight of each criterion should be made.

The definition of clear and evaluable questions is fundamental to a successful evaluation. The number of questions should be limited and it is preferable to address fewer questions with a certain level of rigour and representativeness than to look into a broad range of questions superficially (Morra and Rist, 2009).

**Box. OECD-DAC evaluation criteria and indicative questions (OECD-DAC, 2002)**

Relevance: The extent to which the programme is aligned to the priorities and policies of the Organization, target group(s), donor(s) and other stakeholders.

In evaluating the relevance of a programme it is useful to consider the following indicative questions: To what extent are the objectives of the programme still valid? Are the activities and outputs of the programme consistent with its objectives? Are the programme objectives addressing identified needs of the target group(s)? Do the activities and outputs address the problems identified? Is the programme consistent with the C/5 and C/4 (including the relevant expected results, Strategic Objectives and UNESCO's Global Priorities), the Millennium Development Goals and other international development goals? Does UNESCO have a comparative advantage in designing and implementing this programme?

Efficiency: This concerns the relation between the outputs (qualitative and quantitative) and the inputs. An efficient programme uses the least costly resources to generate its planned outputs. In evaluating the efficiency of a programme it is useful to consider the following indicative questions: Have the outputs been delivered in a timely manner? Could the activities and outputs have been delivered with fewer resources without reducing their quality and quantity? Could more activities and outputs have been delivered with the same resources? Have UNESCO's organizational structure, managerial support and coordination mechanisms adequately performed in function of delivery? What measures were taken to ensure that resources are efficiently used?

Effectiveness: The extent to which a programme attains its objectives. It has to do with the immediate effects of programme outputs on target groups (and beyond).

In evaluating the effectiveness of a programme it is useful to consider the following indicative questions: To what extent were the objectives achieved / are likely to be achieved? What were the major factors influencing the achievement or non-achievement of the objectives? To what extent have beneficiaries been satisfied with the programme? Has the programme been cost-effective, i.e. could the outcomes have been achieved at lower cost through adopting a different approach and/or using alternative delivery mechanisms? Does the programme have effective monitoring mechanisms in place?

Impact: The positive and negative (long-term) changes produced by a programme, directly or indirectly, intended or unintended. This includes the main effects resulting from the programme on institutional, social, economic, environmental and other dimensions. Impact goes beyond the immediate effects (effectiveness) of the programme, covering a much broader range of all

(potential) direct and indirect effects.

In evaluating the impact of a programme it is useful to consider the following indicative questions: What real difference has the programme made to the target groups (and beyond)? What changes can be identified that may (in part) be caused by the programme? What are the major factors influencing the identified changes? To what extent can the identified changes be attributed to the programme? To what extent can the changes be measured? Are there any unintended effects of the programme?

Sustainability: This concerns the issue of whether the benefits of a programme are likely to continue after (UNESCO) support/funding has been withdrawn.

In evaluating the sustainability of a programme it is useful to consider the following indicative questions: To what extent did the benefits of a programme continue after (UNESCO) support/funding ceased? What is the likelihood that the benefits from the programme will be maintained for a reasonable period of time if the programme were to cease? Is the programme supported by local/national institutions and adapted to local/national conditions? Are requirements of local/national ownership satisfied? Do partners have the financial capacity to maintain the programme and/or its outputs/outcomes? Are any (other) potential donors interested in maintaining/ further building upon the programme?

*Note: In the literature there is no consensus on the definitions of impact and effectiveness (and consequently no clear-cut distinction between the two). Both criteria refer to the changes brought about by the programme (the effects). In the definition of the ToR one may opt to combine the two criteria into one set of questions on effects.*

## **METHODOLOGY**

This section should specify in broad terms the methodological building blocks of the evaluation. Some level of detail on methodology may be needed in order for (potential) evaluators already to have an idea on the potential approach(es) to be followed and the resources and technical competencies required to undertake the evaluation. Moreover, it is useful to specify any expectations or requirements on the amount and type of data collection in different locations, including an indicative number of field visits/missions/locations to be included in the evaluation design. Finally, an estimation of the total number of consultancy days/weeks in the ToR constitutes a useful parameter that sets the boundaries within which (potential) evaluators are expected to develop a feasible and high-quality methodological proposal.

Too much detail on methodology in the ToR should be avoided as it may preclude any useful methodological suggestions from evaluators. Moreover, it may be too soon in the evaluation process to make informed choices on the most optimal evaluation design. An inception report to be submitted by the evaluation team at the initial stage of an evaluation assignment is usually the preferred modality for developing and agreeing upon the detailed methodological approach.

## **ROLES AND RESPONSIBILITIES**

In this section, the following elements should be specified:

- Who will manage the evaluation?
- Who will be responsible for logistics?
- Who will undertake (what part of) the data collection, analysis and report writing?
- Who is responsible for quality control?
- Is an evaluation reference group being established? What will be its role?

With respect to logistics, it is important to identify requirements such as materials and office space and who will be responsible for it. The evaluator (evaluation team) will commonly be responsible for logistics: office space, administrative and secretarial support, telecommunications, printing of documentation, travel arrangements and logistics for fieldwork, etc. The evaluator is also responsible for the dissemination of all methodological tools such as surveys, but the UNESCO Sector should facilitate this process to the extent possible by providing contact information such as email addresses of relevant stakeholders and networks to be consulted during the evaluation. The donor and implementing partners might be requested to provide planning documents, mission reports or other relevant documents. Some of these (and possibly other) elements should be explicitly mentioned in the ToR to avoid confusion and clarify any issues of (institutional) accessibility, confidentiality and support during the evaluation process.

## **QUALIFICATIONS EVALUATION TEAM**

In this section it is important to specify:

- Whether the evaluation is expected to be conducted by a single evaluator, two evaluators or a team. In case of the latter, a description of the composition of the team and qualifications of different team members is required.
- The qualifications. The evaluator/team should demonstrate (a) extensive knowledge of, and experience in applying, qualitative and quantitative evaluation methods; (b) a strong record in designing and conducting/leading evaluations; (c) expertise on the subject matter of the programme. Detailed knowledge of the role of the UN and its programming is desirable. Additional qualifications and skills areas that should be specified include: language skills (verbal and written), in-country or regional experience. Evaluation teams ideally should be multicultural and consider gender balance and geographical representation.
- The type of proof needed for verifying the qualifications of the evaluation team. This usually includes curriculum vitae as well as two or three examples of recently completed evaluation reports.

When defining the qualification requirements it should be made clear which of the criteria are considered as mandatory (i.e. where non-compliance leads to disqualification of a proposal) and which are optional (i.e. desirable but not a precondition for qualification).

## **DELIVERABLES AND SCHEDULE**

This section should describe in detail the expected deliverables and the timeline for delivery.

Usually, there are three main deliverables in an evaluation:

- An inception report which contains the intervention logic of the programme (based on desk study), an evaluation plan and a list of reviewed documents. The evaluation plan should contain the proposed data collection methods and data sources to be used for addressing each evaluation question/dimension. It is advisable to recommend the use of an evaluation matrix (which connects questions to data collection methods/sources).
- A draft evaluation report. A stakeholder workshop, facilitated by the evaluation team to discuss preliminary evaluation findings and conclusions may be considered at this stage. The timeline of the evaluation should allow for sufficient time for stakeholders to review and discuss the draft report.
- A final evaluation report.

It is recommendable to specify the structure of the report in the ToR. An example of a common structure is the following:

- Final report (specify the number of pages, excluding Annexes), to be structured as follows:
  - o Executive Summary (usually not more than 3–4 pages)
  - o Programme description

- Evaluation purpose
- Evaluation methodology
- Findings
- Lessons learned
- Recommendations
- Annexes (including the list of stakeholders consulted during the evaluation, key documents reviewed, ToR, survey forms and aggregate findings, budgetary analysis, etc.)

## **REFERENCES**

Any references to key documents available on the internet or distributed along with the ToR should be listed here, as these can provide useful insights and guidance for evaluators in preparing their proposed approach and methodology.

References for this document:

Morra Imas, L.G. and R. Rist (2009) *The Road to Results – Designing and Developing Effective Development Evaluations*, The World Bank, Washington D.C.

OECD-DAC (2002) *Glossary of Key Terms in Evaluation and Results Based Management*, OECD-DAC, Paris.