

# PLANNING AND CONDUCT OF EVALUATIONS

## 6. Planning for Scope and Methodology

### Principles and Norms

#### ENSURING QUALITY OF EVALUATION

6.1 Each evaluation should employ processes that are quality oriented, and use appropriate methodologies for data collection, analysis and interpretation.<sup>37</sup>

Based on UNEG Norm 8

#### COVERAGE OF EVALUATION AND TERMS OF REFERENCE

6.2 Each evaluation should be planned and a TOR drawn up to:

- Define the rationale, purpose, and scope of the evaluation, including the specific objectives of the evaluation and the main audience for its findings. The purpose will almost always include an assessment of performance and results to date, and will also answer other strategic questions relating to governance and management, financing, scope, or a particular policy. (See also Chapters 16 and 17 for complete checklists with respect to TORs and the content of evaluation reports.)
- Define clearly the subject of the evaluation — a GRPP or subset of its activities, or possibly several GRPPs in the same sector — as well as the contextual factors and issues that need to be understood and that affect the methodology chosen. In the case of a GRPP, essential contextual background information includes the circumstances surrounding the origin of the program, its maturity, its objectives, the coverage and range of activities supported, the identification of stakeholders, trends in expenditures, and expected outcomes and impacts for specific target groups.<sup>38</sup>
- Define the criteria by which the program will be judged. In

Elaborates on DAC Principle IX, para. 32

---

37. The U.S. Program Evaluation Standards and the African Evaluation Guidelines have “accuracy standards” related to information gathering, analysis, drawing of conclusions, reporting, and meta-evaluation.

38. This treatment should include the “raison d’être” of the program — that is, why global or regional collective action was deemed necessary or useful, and what additional features the partnership brings to the program.

addition to the standard criteria of relevance, effectiveness, efficiency, and sustainability, the legitimacy and effectiveness of the governance and management arrangements will almost always need to be covered.

- Define the evaluation issues and questions that will be addressed in the evaluation, such as the continued rationale for the program, the achievement of objectives, factors influencing the achievement or non-achievement of objectives (both internal and external to the program), and other outcomes and impacts (both intended and unintended). Aspects of the performance of the host organization and/or the program's partners could also be addressed where the host organization is performing some governance or management functions on behalf of the program and where the partners have made specific commitments to the program (such as pledges to provide funding). However, the inclusion of these in the TOR should be cleared with the host organization and the partners, respectively. (See also paragraph 12.9.)
- If desired, allow some evaluation issues and questions to remain open until key stakeholders have been consulted during the course of the evaluation, in case additional issues are raised from those stakeholders' perspectives. Any substantial changes in the objectives and scope of the evaluation should be communicated to the commissioner of the evaluation, and an amended TOR or budget approved as required.
- Define the methods and techniques to be used to address the issues identified, including proposed methodologies for gathering of information (existing or new), analysis of the information, and interpretation of the results of the analysis. For GRPPs, it is normally necessary to collect data on results at the program, portfolio, and activity levels. It is particularly important to have a representative sample of activities, since GRPPs have a large scope and multiple beneficiaries. An attempt should be made to establish causal relationships in accordance with an existing results chain or logframe, if available, while acknowledging the complexities and identifying assumptions and limitations. If a comparison to a counterfactual is to be attempted, stakeholder consultation is desirable to agree on the counterfactual.<sup>39</sup>
- Determine the resources and time required to complete the evaluation.

---

39. The *counterfactual* is the situation or condition that hypothetically would have prevailed if there had been no development intervention. OECD/DAC, *Glossary of Key Terms in Evaluation and Results Based Management*, 2002.

## Standards and Guidelines

### RATIONALE, PURPOSE, AND OBJECTIVES OF AN EVALUATION

6.3 The *rationale* of the evaluation describes why and for whom the evaluation is undertaken and why it is undertaken at a particular point in time. The *purpose* of the evaluation is usually in line with the learning and accountability functions of evaluations – such as (a) to contribute to improving the effectiveness of the program; (b) to consider a continuation, discontinuation, or change in scope of the program; or (c) to account for aid expenditures to stakeholders and taxpayers. The *objectives* of the evaluation specify what the evaluation aims to achieve – such as (a) to ascertain the results of the program; (b) to assess the relevance, effectiveness, and efficiency of the program; and (c) to provide findings, conclusions, and recommendations with respect to specific aspects of the program. As noted in Chapter 2, both the purpose and objectives of the evaluation are likely to be different at different stages of program maturity. (See Tables 5 and 6 for more detailed guidelines.)

Based on DAC Standards 1.1, 1.2 and 1.3

6.4 The evaluation objectives should follow from the purpose of the evaluation. They should be realistic and achievable – taking into consideration the scope of the program, the number of activities, and the quality of data on the one hand, and the overall time frame, resources, and level of participation and consultation expected on the other hand. The objectives of the evaluation should be clear and agreed upon by all partners involved.

Elaborates on UNEG Standard 3.5

### SCOPE OF THE EVALUATION

6.5 The scope of the evaluation should be clearly articulated by defining the time period covered by the evaluation, the interventions and activities to be included in the evaluation, and any delimitation on geographical or target group coverage. Any restrictions in scope should generally be justified by the rationale and purpose of the evaluation and explained in the evaluation report. In general, restrictions in scope are not justified for evaluations conducted for the purpose of satisfying accountability needs.

Elaborates on DAC Standard 2.1 and UNEG Standard 3.5

6.6 GRPPs are distinguished by the expectation of benefits arising from the partnership, over and above the benefits associated with the discrete activities supported (whether global, regional, country, or local activities). These additional benefits of the partnership may arise from the large scale, from joint activities enabled by the partnership, or from the cross-fertilization and enrichment of knowledge among the large number of partners. Thus, the scope of a GRPP evaluation should ideally encompass the achievement of these additional expected benefits, not just the benefits of the discrete activities supported.

6.7 The purpose and scope of the evaluation are likely to be different at different stages of program maturity. Tables 5 and 6 provide broad guidelines for determining the evaluation scope and questions at different stages of the program.

**Table 5. Sample Issues to Feature in the Scope of an Evaluation at Various Stages of the GRPP**

Stage/Timing	Sample Issues to Be Examined
A. Early Stage (first 2–3 years)	Design of the program: Is it appropriate or in need of adjustment?
	Institutional structures of the program: Are governance and management arrangements in place and functioning as planned?
	Resource mobilization: Have the needed resources been mobilized for governance, management, and M&E needs, and is there a strategy for growth to support a growing portfolio?
	Mechanism for monitoring and evaluation: Have appropriate M&E mechanisms been set up?
	Program performance: Using process indicators, are input, activity, and output targets being met?
	Capacity building: Are steps being taken to close the gaps in the capacity of national or local institutions, as applicable?
B. Established Stage (over 5 years old)	Operations: Are these functioning as designed?
	Sources and uses of funds: Are the inflows of funds stable or growing? Are the allocation processes and reviewing of proposals (if applicable) working as planned?
	Targets: Are these being met?
	Capacity strengthening: Are national institutions being strengthened? Is technology being transferred?
	Outcomes and impacts: Are expected goals being met?
	Strategic direction: Given the above, are the program's strategic directions correct and on course? For growth, outreach to new donors and partners, broadening the target area, devolution, etc.?
C. Mature Stage	Outputs: Are outputs matching planned expectations?
	Impacts: Are there measurable indicators of the impacts of the program?
	Capacity optimization: Are national personnel and institutions capable of performing effectively?
	Sustainability: What measures have been taken to ensure the sustainability of the program with regard to financial, institutional, and other resources?
	Strategic direction, devolution, and possible exit: What arrangements have been made for the planned transfer of implementation responsibilities or withdrawal of external resources — financial, technical, etc.?

*Source:* Both this and the following table have been adapted from material provided by Dr. Adetokunbo Lucas, based on his experience in evaluating GRPPs and in serving on the governing bodies of GRPPs.

**Table 6. Schematic Representation of a Life-Cycle Approach to Determining the Scope of an Evaluation**

Issues	Program Stage		
	Early	Established	Mature
Program Design	+++	++	+
Governance Structures	+++	++	+
Management Structures	+++	++	+
Resource Mobilization Strategy	+++	+++	++
Inputs	+++	+++	++
Activities	+++	+++	+++
Outputs	++	+++	+++
Reach	++	+++	+++
Outcomes	+	++	+++
Impacts	+	+	+++
Sustainability	+	++	+++
Devolution or Exit Strategy	+	++	+++
Miscellaneous Topical Issues	?	?	?

Priority: High +++ , Medium ++ , Low + , To be determined ?

#### FACTORS AFFECTING THE CHOICE OF METHODOLOGY

6.8 The objectives and scope of the evaluation are critical references in determining the evaluation methodology. The issues and questions to be addressed, the type of information sought, the log-frame (if available), the nature of the results chain, and the quality of data already available also affect choice of methodology. Within this broad context, evaluation methodologies should be sufficiently rigorous to ensure a complete, fair, and unbiased assessment of the GRPP. Budgets should be sufficiently flexible to allow for this level of rigor and quality. Another important factor that affects the methodology is the level of participation and consultation of stakeholders desired. The evaluation TOR should include a “stakeholder map” of the various stakeholders and their roles, and indicate the level and type of participation expected. (See paragraph 4.8.)

Elaborates on  
UNEG  
Standards 3.5  
and 3.7

6.9 GRPPs are highly diverse, and evaluation questions and methodologies need to be tailored to the specific sectors in which they operate. Where applicable, evaluation questions should consider whether private sector behavior or commercial market forces have influenced results, and what would be the best way of capturing these effects, taking into account the different incentives involved.

6.10 Care should be taken in planning the methodology to consider up front whether and how issues relating to gender and under-represented groups are to be addressed.

Based on UNEG  
Standard 3.7

## ENSURING AN APPROPRIATE CHOICE OF METHODOLOGY

6.11 Since most GRPPs do not have a specialized evaluation unit or professional evaluators on the governing body, it is advisable to plan for the input of expert evaluators in determining the methodology before the TOR and contracts are finalized. Evaluators from the specialized evaluation units of one or several partners may be called upon to contribute, or consultants may be hired specifically to advise on planning the evaluation (in which case they should not be eligible to compete to conduct the evaluation). Alternatively, a skilled peer group could be called on to advise the governing body and review the TOR to ensure that appropriate parameters of evaluation design are specified.

Draws on UNEG  
Standard 3.10,  
para. 21

6.12 While the evaluation methodology should be planned up front and specified in the TOR, the processes of planning and managing evaluations may also provide for the teams that are bidding to conduct the evaluation to make specific proposals, or for the team, once selected, to provide further details of the methodology in an inception report in the early stage of the evaluation process. Under all circumstances, it is helpful for the commissioner of the evaluation to interact with the evaluation team before work begins in order to clarify expectations. Any amendments to the scope and methodology that result from these processes should be endorsed by the governing body commissioning the evaluation and reflected in a formally revised TOR, which is then attached to the final evaluation report. The evaluation team should be held to the final TOR.

## ABSENCE OF AN ADEQUATE M&E FRAMEWORK

6.13 If the evaluation team finds that the objectives of the program are unclear, that an M&E framework is lacking, or that needed data are absent or of questionable quality, the team may advise that an evaluability assessment be conducted as a prerequisite to the evaluation. (See paragraph 2.7.) Also, if there is an M&E system, but it is not producing the quality of information necessary for an evaluation, a process audit might be advisable to see what constraints have been limiting its usefulness. In either case, there is a need to define a time frame for a decision (a) to create the conditions for a successful evaluation or (b) to go ahead with an evaluation, while acknowledging the inevitable limitations.

6.14 If it is decided to go ahead and evaluate with limitations, the evaluation team could agree with the commissioner of the evaluation to reconstruct a results framework and baseline information, as needed. This may require changing the TOR, the time frame, or the budget for the evaluation. (See also paragraphs 9.5 and 9.6.)

6.15 Key actions to create enabling conditions for quality evaluations are outlined in Chapter 2. These include (a) establishing or adapting the M&E framework to meet the requirements for evalua-

tion of the GRPP (such as collecting information on all interventions at all levels); (b) assigning clear responsibilities for M&E; and (c) ensuring that the responsibilities for monitoring and evaluation are separate in order to assure independence of evaluation.

#### USE OF EXISTING EVALUATIVE INFORMATION

6.16 Ideally, the evaluation of the performance and results of a GRPP relies not only on information at the program level, but also on summary portfolio information and on more detailed information at the country and activity levels in order to validate overall findings. When using information from the country or activity level, evaluators should always make explicit whether they have merely accepted information or ratings provided by management or by the country (which would generally be only self-evaluations), or whether they have undertaken to validate such information through independent assessments based on field observations.

Draws on IEG's experience with reviewing GRPPs

#### EVALUATION CRITERIA

6.17 The most common criteria for evaluating development assistance, endorsed by both the DAC and the UNEG, apply equally to GRPPs. These are relevance, effectiveness (or efficacy), efficiency, impact, and sustainability. The UNEG Standards also list additional criteria that may be used: value-for-money (which is an aspect of efficiency) and target group satisfaction.<sup>40</sup>

Based on DAC Standard 2.3 and UNEG Standard 3.6

6.18 To aid the reader, the criteria should be defined in unambiguous terms in the report. If one of the usual criteria specified above is not applied, this should be explained in the evaluation report, along with any additional criteria that were used. (See Chapters 9 to 15 for more guidance on evaluation criteria and questions relevant to GRPPs.)

Based on DAC Standard 2.3

#### CONSIDERING POSSIBILITY OF PEER REVIEW

6.19 Depending on the scope and complexity of the evaluation, it may be useful to establish a peer review or reference group composed of experts in the technical topics covered by the evaluation, or experts in evaluation itself. This group would provide substantive guidance to the evaluation process (such as providing feedback on the methodology, analysis, and interpretation of results) and provide quality control of the draft reports.

Based on UNEG Standard 3.12

---

40. One multilateral agency uses the criterion of additionality or value added. There may be trade-offs in achieving results relating to different criteria (such as between effectiveness and efficiency), which may need to be taken into account in assessing the performance of the program in relation to each criterion. (See also Chapter 12, Governance and Management.)