EVALUATION GOVERNANCE ISSUES

2. Prerequisites and Enabling Conditions for Effective Evaluations

Principles and Norms

INSTITUTIONAL RESPONSIBILITY FOR MONITORING AND EVALUATION

2.1 Appropriate institutional arrangements for managing M&E are a prerequisite for ensuring effective processes and for making full use of the information generated by M&E systems. Plans for proper monitoring and periodic evaluation should be built into the design of the program at inception. Institutional arrangements need to meet the requirements for (a) a policy and set of guidelines for M&E, including a disclosure policy; (b) impartiality and independence of evaluations; and (c) using M&E findings to improve future decision making and activities.

2.2 Monitoring is always the responsibility of the management and operational staff, and evaluation is the responsibility of the governing body or other unit separate from management. In most GRPPs, evaluations are commissioned by part-time governing bodies and conducted by independent teams of consultants or independent experts. In larger GRPPs, there may be a mandate and sufficient resources for a separate internal evaluation unit.11 In either case, the body commissioning the evaluation takes responsibility for the quality of the final report and for disseminating the findings and recommendations, in different formats for different audiences, as appropriate.12

MONITORING AND EVALUATION POLICY

2.3 Existing principles and norms issued by the DAC and the UNEG stress the need for a policy on evaluation approved by the

11. Where there is a separate evaluation unit, the principle on competency should also apply: “Aid agencies/ [programs] need a critical mass of professional evaluation staff in order to have sufficient expertise in their various fields of activity and to ensure credibility of the process.” (DAC Principle IV, para. 19.)

12. Having a governing body commission or manage an evaluation that includes an assessment of governance (that is, the performance of the governing body itself) poses a potential conflict of interest. In some cases, this may be resolved by designating a “higher body” or “external group” to which the evaluation team would report directly. See also paragraphs 3.5 and 3.6 on organizational independence and Chapter 12, Governance and Management.
governing body. At a minimum, the governing body should make an early commitment in the policy to the following:

- To have the GRPP evaluated periodically and to provide adequate funding for conducting evaluations
- To agree on the purpose of regular evaluations
- To agree on how evaluation results will be disseminated and used to improve accountability, learning, decision making, and broader knowledge sharing (including among evaluators).

2.4 The policy should also address such issues as independence and impartiality of evaluation, desired stakeholder participation and consultation, and the openness of the evaluation process (including disclosure). The policy should define the respective roles of management and governing bodies in M&E, and explain how evaluations are expected to be planned, managed, budgeted, and reviewed. The policy should also include guidance on mandatory criteria and on processes for selecting external evaluation teams, if applicable. (See the standards section below for further guidance.)

2.5 The monitoring and evaluation policy should also refer to the need for an effective monitoring system, both to provide the information required for scheduled reporting to the governing body on the use of resources, the progress of activities, outputs, and outcomes, and to provide the information necessary for future evaluations.

**MONITORING AND EVALUATION FRAMEWORK**

2.6 A key enabling condition for effective evaluation is the early establishment of an M&E framework. Early after its launch, each GRPP should put in place an M&E framework, at least at the program level, which includes (a) clear and coherent objectives and strategies, (b) an expected results chain, (c) measurable indicators that meet the monitoring and reporting needs of the governing body and management, and (d) systematic and regular processes for collecting and managing data, including baseline data. (See standards and guidelines below for establishing an effective M&E framework.) An approximate date for the first independent evaluation should be set (generally 2–3 years after the launch of the program), and the evaluation budgeted.

2.7 An evaluation may have to include — or be preceded by — an assessment of the adequacy of the M&E framework and system, if there are doubts about their adequacy. Such an evaluability assess-
ment would determine whether the objectives of the GRPP have been clearly defined; whether *SMART* indicators have facilitated the collection of timely, relevant, and accurate data; whether information sources are accessible and reliable; and whether any serious constraints prevent an impartial evaluation process. The commissioners of the evaluation should decide whether the evaluability assessment should precede or be part of the evaluation itself.

**MONITORING AND EVALUATION PLANNING AND PROGRAMMING**

2.8 M&E requirements for reporting, accountability, and learning need to be built into the regular planning processes in GRPPs from the start. (See standards and guidelines below for more information on establishing monitoring systems.)

2.9 Planning and programming of evaluations must take into account the needs of the governing bodies and program managers, as well as those of other potential users of evaluation products (such as policy makers and activity implementers, whether public or private). Otherwise, evaluations should be timed to provide effective input into key decisions, such as approving a new phase, expansion, funding replenishments, restructuring governance, reaching out to new donors, and the like. Other important factors that influence timing include whether the program has reached a certain degree of maturity or completed a critical mass of supported activities. (See also paragraph 6.7.) Cluster evaluations that include other programs with similar objectives, or those operating in the same sector, may be explored. The evaluation plan should identify specific possibilities for interaction with stakeholders and for the participation of various groups of stakeholders. It should also include a dissemination plan.

2.10 Ideally, an evaluation plan calling for broad stakeholder participation would meet all the accountability needs of donors and other partners and would obviate the need for individual stakeholders to undertake separate evaluations. In practice, however, some individual donors may need to conduct separate or joint evaluations with other donors in order to meet their reporting obligations to their authorities, thereby giving rise to multiple evaluations. In such cases, donors and other partners should, at a minimum, share their plans for evaluations and schedule them to facilitate potential complementarity and appropriate programming, since joint, or at least coordinated evaluations

14. For example, the World Bank’s Development Grant Facility (DGF) requires that an independent evaluation be conducted every three-to-five years for the programs that it supports.
tend to achieve greater benefits than separate evaluations in terms of efficiency, consensus-building, and joint learning.  

2.11 The M&E plan, including provisions for any single-donor or joint evaluations, requires the support and endorsement of the governing body. Before any evaluation commences, the full governing body should approve the management and reporting arrangements and the terms of reference (TOR).

RESOURCES AND BUDGETING

2.12 The governing body must allocate adequate resources for the effective implementation and operation of the M&E plan. Planning for monitoring and evaluation should be an explicit part of planning and budgeting for the program as a whole. This includes allocating staff and budget resources to establish feedback mechanisms in order to ensure that the results of evaluations are utilized in future policy and program development.

QUALITY CONTROL

2.13 Quality control must be exercised throughout the evaluation process. Depending on the scope and complexity of the evaluation, quality control is carried out either internally or through an external body, peer review, or reference group. (See also paragraph 2.30.)

Standards and Guidelines

ADDITIONAL FUNCTIONS OF THE INSTITUTIONAL ARRANGEMENTS FOR MONITORING AND EVALUATION

2.14 Ideally, the institutional arrangements for monitoring and evaluation should also (a) promote a culture that values M&E as a basis for learning and improving the effectiveness of the program; (b) provide adequate financial resources for M&E; (c) ensure that accountability for M&E and its follow-up is clear; and (d) enable capac-

---


16. The U.S. Program Evaluation Standards and the African Evaluation Guidelines cite among “feasibility standards” the importance of assessing the “political viability” of the evaluation up front: “The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.”
ity strengthening in M&E, cooperation, and shared learning with other organizations or programs.

**SPECIAL ARRANGEMENTS FOR DRAWING UP THE M&E POLICY AND EVALUATION PLAN**

2.15 Given the large size of some governing bodies, and the possibility that members may not have evaluation experience or expertise, it may be necessary to set up a subcommittee to draw up the evaluation policy and M&E plan. This subcommittee may in turn wish to seek help from an external peer group of representatives of various stakeholders or from expert consultants. The full governing body should approve the final M&E policy and evaluation plan. The use of an external peer group may also be appropriate for overseeing and managing especially complex or difficult evaluations.

**PARTICIPATORY M&E AND EARLY IDENTIFICATION OF STAKEHOLDERS**

2.16 The M&E policy and evaluation plan should give serious consideration to participatory methodologies. The very process of developing a participatory monitoring system tends to enable and inform the planning of participatory evaluations. It also contributes to capacity strengthening and can lay the foundation for sharing experience across activities supported by the program. This, in turn, will enhance program improvement and subsequent evaluation. A high degree of stakeholder participation in prior evaluations is also likely to enhance the quality of subsequent evaluations.

2.17 A comprehensive list of stakeholders, or “stakeholder map,” which also includes information on their perceived roles and responsibilities, is an indispensable prerequisite for determining the participation and consultation process to be followed in conducting evaluations. This map should be updated periodically, and included in the TOR for each evaluation, along with the desired participation and consultation process, to guide the evaluation team. (See also paragraphs 4.7–4.9, and 4.17–4.20.)

**MONITORING AND EVALUATION POLICY**

2.18 The approved GRPP evaluation policy should include:

- A clear explanation of the concept and role of evaluation within the organization
- A clear definition of the roles and responsibilities of the governing body, senior management, task team leaders, and evaluation professionals, if applicable
- Clear guidelines on the process of planning for evaluations
- Clear guidelines on how evaluations are organized, managed, and budgeted

*Based on UNEG Standard 1.22*
• Clear guidelines on the follow-up of evaluations
• A clear statement on and guidelines for disclosure and dissemination.

2.19 The evaluation policy, and related policies on risk management and audit plans, should also be agreed with the host organization in which the program is located, if applicable.

EVALUATION PLANNING AND PROGRAMMING

2.20 Evaluation planning and programming needs to take into account the maturity of the program, since its maturity will likely affect both the purpose of the evaluation and its scope and methodology. The following may serve as a general guideline:

• **Program in Early Stages (first 2–3 years):** Important purposes of an early — usually the initial — evaluation would be to assess the appropriateness of the program design and to review the governance and management arrangements. The evaluation should also review the relevance and clarity of the objectives, identify constraints that make achievement of specific objectives difficult or impossible, and recommend adjustments if necessary.

• **Established Program (over 5 years old):** The evaluation should address inputs, the progress of activities, and outputs. The recommendations will typically focus on ways to increase the effectiveness and efficiency of the program.

• **Mature Program:** At this stage, the program should be operating smoothly and meeting the expectations expressed at its initiation. The evaluation will typically pay particular attention not only to outputs, but also to outcomes, as well as to sustainability and other strategic issues such as growth, devolution, or exit.

2.21 It may be advantageous to undertake evaluations for established and mature programs in the form of “cluster evaluations” done jointly with other programs operating in the same sector. A simultaneous assessment of their objectives, strategies, and activities may identify a potential for achieving joint results more effectively. Complementary or mutually reinforcing inputs, activities, and outputs may also be identified, yielding recommendations for improving collaboration and reducing duplication.

17. See also Chapter 6, Planning for Scope and Methodology.
DESIGN OF THE M&E FRAMEWORK

2.22 The design of an M&E framework at the activity level will vary by program, in terms of both detail and breadth of applicability. Ratings of progress and performance may or may not be used. For most GRPPs, particularly those that do not involve investments, it may not be cost-effective to require an M&E plan or independent evaluation for each separate activity. Rather, for most GRPPs, broad indicators will typically be drawn up for monitoring all similar activities. A subset of activities, defined by size or type, may require periodic progress reports and a completion report at the end. For most programs these will be self-evaluations. These reports will help provide essential information not only for monitoring, but also for evaluations as well.

2.23 If the governing body and/or management decide that an impact evaluation (of longer-term results) will be required in the future, it is important to plan for this from the outset, so that appropriate indicators (possibly for control groups as well as beneficiary groups), data-gathering, quality control mechanisms, and recording of baseline data can be funded and implemented. It may be advisable to contract with an expert evaluation team for the planning of the future impact evaluation, including the collection of baseline data. (See Chapter 15, Impact Evaluation.)

INITIAL STEPS IN ESTABLISHING AN M&E FRAMEWORK

2.24 At the inception of the program, when preparing the budget for the early years, the development of an M&E framework should be adequately funded and provision made for staffing or engaging consultants for its planning and implementation.

2.25 It may be advisable to carry out a “readiness assessment” before the establishment of the system. This is a diagnostic tool for assessing the organizational capacity of the program and the political willingness of its governing body to monitor and evaluate the achievement of the program’s goals and to develop a performance-based framework. The three main objectives of a readiness assessment are (a) to determine whether incentives are appropriate for the suc-

18. The Global Environment Facility has such requirements for its investment grants, since it views the benefits as exceeding costs, and has a critical mass of staff available to carry out the requirements.

19. This section draws on the book by Jody Zall Kusek and Ray C. Rist, 2004, Ten Steps to a Results-Based Monitoring and Evaluation System, as applicable to GRPPs.
cessful launch of an M&E system; (b) to analyze roles and responsibilities in order to be able to assign accountability for the functioning of the M&E system; and (c) to assess existing capacity.

2.26 At a minimum, and to provide an adequate foundation for accountability, the monitoring of inputs (such as expenditures and staff time) should begin immediately after the launch of the program. A classification scheme for expenditures needs to be drawn up after careful discussion. It should differentiate between (a) administrative and activity expenditures; (b) expenditures at the country, regional, and global levels; and (c) other relevant distinctions (such as categories of expenditure, or expenditures by different agencies, and the like) A decision needs to be made whether to record and/or track non-financial contributions/expenditures.

2.27 The next step toward developing an M&E framework is agreement on the objectives of the program. For many programs, consensus objectives will have been defined at meetings preceding the funding of the program and incorporated in the program charter. For other programs, the funding follows agreement on the need to address a particular issue, but the proper response to it, and thus the objectives of the program, are not agreed until later. It may be necessary not only to list objectives, but also to establish an “objectives hierarchy,” should trade-offs emerge later in implementation.

2.28 Once objectives are agreed, a strategy and a set of interventions or activities are agreed upon, generally with the expectation that these will generate specified outputs and outcomes. The early establishment of such an expected results chain is an important step in building an M&E framework. Also important is the early identification of stakeholders and categories of beneficiaries to enable monitoring of welfare outcomes.

2.29 Next, additional indicators that measure the progress of activities, outputs, outcomes, and impacts, and any relevant external factors that affect results must be defined and provision made for their measurement. Other indicators that might be required to assess the achievement of the objectives — such as the success of participatory measures, or the degree in which the program is successful in responding to opportunities and in learning from experience — also need to be defined. Indicators of participation should cover not just participation rates (inputs) but also the expected outcomes such as learning, awareness, behavioral change, and so on. The measurement of the agreed indicators could be the responsibility of activity-level staff in the course of implementation, it could rely on government statistics, or it could be done periodically by specialized program staff (for example, by means of surveys of target group satisfaction).
2.30 Accountability for data collection, monitoring, and reporting needs to be clear, and a quality-assurance system put in place. Once this is done, plans should be made to collect baseline data and ensure that these are stored in an accessible place.

2.31 As the program evolves, strategic plans prepared by the program management and approved by the governing body may define specific expected outcomes and implementation targets. Management should also agree with the governing body on a desirable periodicity of reporting, which allows regular tracking of progress against targets and expected outcomes.

2.32 Finally, feedback mechanisms are needed to ensure that the comments of management, the governing body, implementing staff, and data-gatherers themselves are taken into account to help continuously improve the M&E system. The system may also need to be adapted periodically in response to changing objectives and strategies, or changes in the external environment.

USE OF A LOGICAL FRAMEWORK

2.33 A logical framework, or logframe, is a management technique that is used to develop the overall design of a development project, to improve implementation monitoring, and to strengthen evaluation, by presenting the essential elements of the project clearly and succinctly throughout its cycle. It is a “cause and effect” model that has been widely used by the bilateral and multilateral donor community since the 1970s in order to establish clear objectives and strategies based on a results chain, to build commitment and ownership among the stakeholders during the preparation of a project, and to relate the project’s interventions to their intended outcomes and impacts for beneficiaries.20

2.34 Developing a logframe is as appropriate for GRPPs that have no fixed end-point as it is for projects with an expiration date, as long as the logframe is kept updated and relevant, and incorporates input from new stakeholders as they become involved in the program. But developing a logframe for a GRPP may be more complex because of the larger number of stakeholders and the different levels of program components and activities. Ideally, a single consensus logframe would be developed for the overall program, which would form the basis for the implementation and monitoring of the program and subsequent program-level evaluations. It may also be desirable to supplement this with more narrowly focused logframes for different stakeholder groups, who may have different objectives and who may

---

wish to add additional indicators that reflect their particular interests in the program.

**MONITORING SYSTEMS AND INDICATORS**

2.35 In planning and adopting monitoring systems, GRPPs should use indicators that are SMART — specific, measurable, attainable, relevant, and time-bound.21

**PEER REVIEW OR REFERENCE GROUP**

2.36 It may be useful to establish a peer review or reference group composed of technical experts in the sector concerned and/or M&E experts. This group would provide substantive guidance to the monitoring and/or evaluation process (such as providing inputs on the TOR and providing quality control of draft reports).

**STAKEHOLDER STEERING OR LEARNING GROUP**

2.37 “When feasible, a core learning group or steering group composed of representatives of the various stakeholders in the evaluation may be created. This group’s role would be to act as a sounding board, and to facilitate and review the work of the evaluation. In addition, this group may be tasked with facilitating the dissemination and application of the results and other follow-up action,” particularly with their own constituents.

**KNOWLEDGE MANAGEMENT AND DISSEMINATION**

2.38 GRPPs should support knowledge sharing by ensuring the highest standards in accessibility and presentation of M&E products, by using a range of channels to reach target audiences, by participating in knowledge management, and by sharing activities with other relevant organizations.

---

21. In the case of indicators, the GEF Policy also adds “realistic” for R, and “timely, traceable, and targeted” for T. The book, *Ten Steps to a Results-Based Monitoring and Evaluation System* also cites useful criteria under the acronym “CREAM”: Indicators should be **Clear** (precise and unambiguous); **Relevant** (appropriate to the subject at hand); **Economic** (available at a reasonable cost); **Adequate** (providing a sufficient basis to assess performance); and **Monitorable** (amenable to independent validation). See also Salvatore Schiavo-Campo, 1999, “‘Performance’ in the Public Sector,” p. 85.