



MANAGING EVALUATIONS: A HOW-TO GUIDE FOR MANAGERS AND COMMISSIONERS OF EVALUATION





Managing Evaluations: A How-To Guide For Managers and Commissioners of Evaluations

www.worldbank.org/ieg/ecd

© 2015 Independent Evaluation Group, The World Bank Group
1818 H St., NW
Washington, DC 20433
<http://ieg.worldbankgroup.org>

IEG: Improving the World Bank Group's Development Results through Excellence in Evaluation

The Independent Evaluation Group is an independent unit within the World Bank Group; it reports directly to the Bank's Board of Executive Directors. IEG assesses what works, and what does not; how a borrower plans to run and maintain a project; and the lasting contribution of the Bank to a country's overall development. The goals of evaluation are to learn from experience, to provide an objective basis for assessing the results of the Bank's work, and to provide accountability in the achievement of its objectives. It also improves Bank work by identifying and disseminating the lessons learned from experience and by framing recommendations drawn from evaluation findings.

IEG's Evaluation Capacity Development Working Papers are an informal series to disseminate the findings of work in progress to encourage the exchange of ideas about development effectiveness through evaluation.

The findings, interpretations, and conclusions expressed here are those of the author(s) and do not necessarily reflect the views of the Board of Executive Directors of the World Bank or the governments they represent, or IEG management.

IEG cannot guarantee the accuracy of the data included in this work. The boundaries, colors, denominations, and other information shown on any map in this work do not imply on the part of the World Bank any judgment of the legal status of any territory or the endorsement or acceptance of such boundaries.

ISBN-13: 978-1-60244-258-0

ISBN-10: 1-60244-258-4

Contact: IEG Communication, Learning, and Strategies

e-mail: ieg@worldbank.org

Telephone: 202-458-4497

Facsimile: 202-522-3125

<http://ieg.worldbankgroup.org>

Acknowledgments

This how-to guide was written by Arianne Wessal, Nidhi Khattri, and Dawn Roberts. It draws on materials developed by Penny Hawkins (Head Evaluation, United Kingdom, Department for International Development) for a workshop conducted at the International Program for Development Evaluation Training as well as a number of other sources. Peer review comments from Clay Wescott and Anthony Martin Tyrell (IEG) are gratefully acknowledged.

The IEG Blue Booklet series disseminates practical information on various aspects of evaluation. An objective of the series is to focus on implementation of evaluation and to disseminate information on other sources of knowledge.

The series is prepared under the team leadership of Nidhi Khattri, Lead Evaluation Officer, and the overall guidance of Monika Weber-Fahr, Senior Manager, IEGCS.

Abbreviations

AEA	American Evaluation Association
DAC	Development Assistance Committee, Organisation for Economic Co-operation and Development
DPME	Department of Monitoring and Evaluation (Republic of South Africa)
IEG	Independent Evaluation Group
NEPF	National Evaluation Policy Framework (Republic of South Africa)
OECD	Organisation for Economic Co-operation and Development
RFP	Request for proposal
ToR	Terms of reference

Table of Contents

Acknowledgments	3
Abbreviations.....	4
Introduction.....	7
Stage 1: Planning the Evaluation	9
Step 1: Consult the Commissioning Entity’s Evaluation Policy	11
Step 2: Define the Purpose and Feasibility of the Evaluation.....	13
Step 3: Prepare a Work Plan to Manage the Evaluation	18
Stage 2: Commissioning the Evaluation.....	20
Step 1: Prepare the Evaluation Terms of Reference.....	20
Step 2: Prepare the Request for Proposals and Select the Evaluator/Evaluation Team.....	23
Stage 3: Undertaking the Evaluation	26
Step 1: Supervise the Evaluation Work Plan and Implementation.....	26
Step 2: Ensure Evaluation Quality	29
Step 3: Provide Feedback on the Inception, Draft, and Final Reports.....	31
Stage 4: Reporting and Follow-Up	33
Step 1: Communicate Results	33
Step 2: Follow Up on Use of Evaluation Recommendations	34
Step 3: Wrap Up the Evaluation	35
Appendix A: Sample National Evaluation Policies.....	37

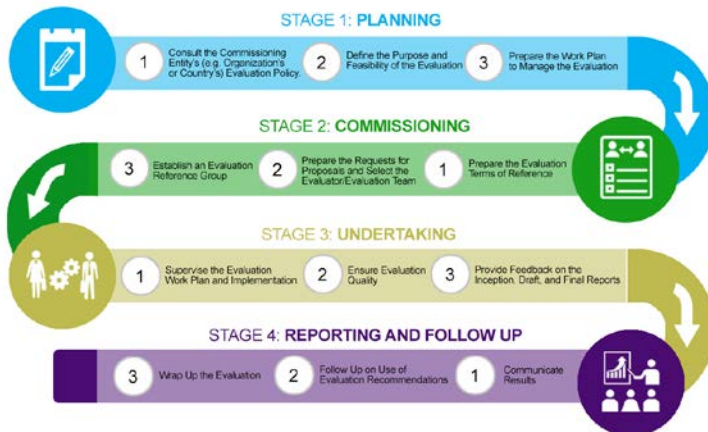
Appendix B: A Guide to Developing Good Recommendations in Evaluation Reports	40
Appendix C: Examples of Follow-Up Mechanisms for Tracking the Adoption of Recommendations	41
Bibliography	48

Introduction

Evaluations serve two main functions—accountability and learning. They provide an “account” to stakeholders about the extent to which resources are being, or were, used efficiently and effectively, and degree to which the intended results are being, or were, achieved. In addition, they are often intended to generate empirical knowledge about the interventions in pursuit of given objectives: what elements of the intervention worked, what did not work, and why. By contributing to such knowledge, evaluations help policy makers, program managers, and other stakeholders make informed decisions regarding policy priorities, program design and implementation, and resource allocations.

This guide provides simple and practical advice for the management and commissioning of evaluations in a public sector context. It provides an overview of the key concepts and processes and covers the major stages involved in managing an evaluation (see Figure 1).

Figure 1: Managing an Evaluation: Major Stages

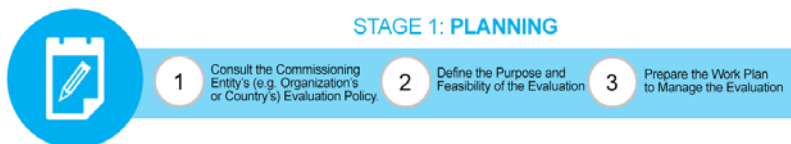


Managing an evaluation can be challenging. Technical issues can become politicized, and evaluators sometimes are viewed with suspicion and face resistance, a lack of understanding, and frustration from local stakeholders, including program implementers and program beneficiaries. Evaluations may also be difficult to undertake because of contextual challenges such as limited

availability of data, weak program management systems, politically unstable environments, corruption, instability and conflict, low human resource capacity, and low levels of transparency and access to information. These issues can affect an evaluation in any setting but may be particularly pronounced in a developing country context. The general stages and steps in this guide should therefore be supplemented with local knowledge and expertise where possible for the most effective management of evaluation.

The terms “evaluation commissioner” and “manager” used in this document refer to any person(s) whose primary responsibility it is to oversee and coordinate the evaluation. Different phases of the evaluation and associated tasks may well be managed by different people. However, these people are likely to liaise with the person(s) with the primary responsibility for overseeing the evaluation.

Stage 1: Planning the Evaluation



A successful evaluation is based on good planning. Good planning helps ensure that the evaluation runs smoothly and minimizes the number of adjustments needed once it has begun. As part of the planning process, it is useful to cover the following basic steps:

- Step 1: Consult the commissioning entity's (for example, organization's or country's) evaluation policy.
- Step 2: Define the purpose and feasibility of the evaluation.
- Step 3: Prepare the work plan to manage the evaluation.

Planning an evaluation is often an iterative process, as information that one thought was available may not be available, or the methods envisaged to carry out the evaluation may turn out not to be appropriate. Therefore, the planning should continue until all the details have been worked out.

Box 1 describes two main types of evaluation to help set the stage for planning.

Box 1: Common Types of Evaluation

Below is an indicative list of common types of evaluation.

Formative evaluations assess ways in which a program, policy, or project can be or is implemented. They are undertaken before or while the program is under implementation, and the focus is typically on design or improvement of implementation.

Summative (or ex post) evaluations, in contrast, focus typically on the outcomes of a project, policy, or program, and to a lesser degree on how implementation could have been improved.

Formative evaluation includes several types of evaluation and assessments:

- **Needs assessment**—Determines who needs the program, how great the need is that the program would address, and what specific intervention(s) might work to meet the need
- **Prospective evaluation**—Helps predict the success or failure of a program by reviewing completed evaluations of similar programs to identify lessons and issues and by looking for ways to strengthen the proposed intervention(s), if the program goes forward
- **Process evaluation**—investigates the process of delivering the program – the intervention(s) – and is geared to answering several types of questions – the quality of management, the relevance of the intervention to the goals of the program, the efficiency of the intervention, and so forth.

Summative evaluation can be subdivided as well:

- **Outcome evaluations** investigate whether, and the degree to which, the program outcomes were achieved. They often disaggregate the outcomes by different demographic groups of interest. They can also look at both intended and unintended outcomes of the intervention.
- **Impact evaluations** are intended to assess the degree to which an intervention caused the outcome of interest, and they often use comparative methods to establish the causal link between the intervention and the outcome.
- **Cost-effectiveness and cost-benefit analysis** address questions of efficiency by analyzing outcomes in terms of their dollar costs and values.
- **Meta-analysis** integrates the outcome estimates from multiple studies to arrive at an overall conclusion on an evaluation question.

Step 1: Consult the Commissioning Entity's Evaluation Policy

Many government and nongovernmental organizations have evaluation policies. In the public sector, evaluation policies provide a framework for processes, structures, and principles to guide a program of, as well as individual, evaluations. Evaluation policies differ by organization but often include information such as when evaluations will be conducted, how often, and under what circumstances; they may also include guidelines on quality standards and ethical norms for carrying out evaluations. Within a government, there may be varying evaluation policy guidelines at the national, subnational, and ministry/department levels. Box 2 provides an example of a national evaluation policy and its key elements; other national evaluation policy examples are provided in Appendix A for comparative purposes.

Evaluation policies at the country level typically set out the country's plan for conducting evaluations. These policies provide details on the objectives the country expects to achieve by carrying out evaluations. They also set out the mechanisms for managing and coordinating evaluations across the whole of government. Government evaluation policies will typically describe steps for linking evaluation with other policy processes at different levels of government, such as planning and budgeting functions. Usually governments will have differentiated guidelines for the national, subnational, and organizational levels, which include guidance on evaluation activities to be carried out over the short, medium, and longer terms.

Box 2: A National Evaluation Policy—South Africa’s National Evaluation Policy Framework

South Africa has a formally written and mandated National Evaluation Policy Framework (NEPF), which was approved by its Cabinet in 2005 and developed and finalized in 2011. Below are key elements of the framework.

The NEPF provides the basis for a system of evaluation across government. Its main purpose is to promote quality evaluations that can be used for learning to improve the effectiveness and impact of government, by reflecting on what is working and what is not working and revising interventions accordingly.

Purpose: To ensure that credible and objective evidence from evaluation is used in planning, budgeting, organizational improvement, and policy review, as well as ongoing program and project management, to improve performance.

Definition of Evaluation: The systematic collection and objective analysis of evidence on public policies, programs, projects, functions and organizations to assess issues such as relevance, performance (effectiveness and efficiency), value for money, impact, and sustainability and to recommend ways forward.

Evaluation Types/Approaches: Diagnosis, design evaluation, implementation evaluation, impact evaluation, economic evaluation and evaluation synthesis. These evaluations can occur at different stages – prior to an intervention, during implementation, and after implementation.

The seven key elements of the NEPF are:

1. Large or strategic programs or those of significant public interest or of concern must be **evaluated at least every five years**. The focus is on government’s priority areas: health, crime, jobs, rural development, and education.
2. Rolling three-year and annual **national and provincial evaluation plans** must be developed and approved by the Cabinet and Provincial Executive Councils. These will be developed by the Department of Monitoring and Evaluation (DPME) and the Offices of the Premier. These plans identify the minimum evaluations to be carried out – departments are free to carry out additional evaluations.
3. The results of all evaluations in the evaluation plan must be in the **public domain**, on departmental and DPME websites (excluding classified information).
4. **Improvement plans** to address the recommendations from the evaluations must be produced by departments, and their

implementation must then be monitored.

5. Departments are responsible for carrying out evaluations. DPME and (in time) the Offices of the Premier provide **technical support and quality control** for evaluations in the national and provincial evaluation plans.
6. Appropriate **training courses** are provided by Public Administration Leadership and Management Academy (now the National School of Government), universities, and the private sector to build evaluation capacity in the country.
7. DPME will produce a series of **guidelines and practice notes** on the detailed implementation of the policy framework, elaborate various aspects of the system, and set quality standards for evaluations.

Source: Policy Framework for the Government-Wide Monitoring and Evaluation Framework, Republic of South Africa 2011.

Jointly conducted evaluations, between governments or organizations, can increase the legitimacy of findings and reduce the overall number of evaluations; however, they can also bring unique challenges when decisions are required to align with multiple evaluation policies. For this situation, the Organisation for Economic Co-operation and Development's (OECD) *Guidance for Managing Joint Evaluations* (2006) provides practical advice for increasing the effectiveness of evaluation work being commissioned by more than one agency (in the case of evaluations initiated or commissioned by development agencies).

Finally, a formal policy framework to guide evaluations is part of strengthening an evaluation system and culture and therefore the use of evaluation results. Evaluation commissioners or managers in countries or agencies that have an evaluation policy will begin by articulating how any given evaluation fits with the evaluation policy and then work to better define the purpose and rationale for the evaluation (Step 2). Evaluation commissioners or managers in countries and agencies without a formal policy might find it useful to invest in an extra up-front effort to explore with key stakeholders how the evaluation will contribute to systematic learning and use of evaluation findings and recommendations.

Step 2: Define the Purpose and Feasibility of the Evaluation

The reasons for conducting an evaluation should be clearly defined before a work plan for its management can be developed. Evaluation findings are most likely to be useful when the evaluation manager determines the main rationale and purpose(s) for the

evaluation, together with intended users in consultation with relevant stakeholders, managers can develop a statement describing the purpose of the evaluation by answering the following questions (CDC 2011a):

- What does this evaluation strive to achieve?
- Who will use the findings?
- How will the findings be used?

The process of developing the initial statement tends to be iterative, as the evaluation manager(s) seek to specify why the evaluation is being conducted, who the main audiences are, how the findings are likely to be used, and what is feasible for the scope given resource and time considerations. Relevant stakeholders should therefore be involved early in defining the rationale and purpose of the evaluation by identifying issues to be addressed and shaping the evaluation questions to be answered.

To ensure that key groups and perspectives are included at appropriate points throughout the evaluation, it is important for the evaluation manager to consider the following questions:

- Who are the stakeholders for this evaluation?
- What is their role in the evaluation?
- What do they need to learn from the evaluation?
- How will the findings and recommendations be communicated?
- How will the findings and recommendations be used?
- What do intended users view as credible information?
- How will the evaluation findings be communicated?

Stakeholders might include:

- Beneficiaries—Those persons, groups, and organizations, who directly benefit from the intervention
- Indirect beneficiaries—Those persons, groups, and organizations who have no direct contact with an intervention but who benefit from it/ or are somehow affected by it nonetheless
- Donors—The organizations or individuals that provide funding for the project or program
- Staff and volunteers
- Politicians (including opposition party members), community members, and other interest groups
- Government officials—Elected officials, government employees with

- common or relevant interest
- Social partner organizations (employer and trade union representatives, representatives of nongovernmental organizations and other third-party organizations)
- Research and advocacy groups.

Stakeholders may participate in focus groups to discuss the type, scope, and criteria for the evaluation; to provide feedback on evaluation questions and key issues; to explore the evaluation methodology, terms of reference, work plan, and data collection strategy and process; as well as to disseminate and gather feedback on results.

Engaging key stakeholders early in the planning stage will provide opportunities to consider whether there are any conflicts between the interests of the evaluation managers and the stakeholders and how these can be addressed. It is important to resolve any disagreements or misunderstandings about the overall purpose or specific evaluation questions among stakeholders. This consultative process will help foster the buy-in or shared ownership that will be valuable in later stages of managing the evaluation.

The process of clarifying the rationale for an evaluation often includes having key stakeholders agree on the main evaluation questions. The selection of one or more common reasons for conducting an evaluation can help a manager and intended users identify the main focus and how the findings will be used.

Evaluations are typically undertaken for learning or accountability purposes. In that regard they may:

- *Contribute to improving a specific development policy, program, or project.* Evaluation supports learning—to determine what approaches and processes worked well or did not work well for improving current and future outcomes. Drawing on monitoring and additional data, evaluations provide managers and stakeholders with a means to reflect on performance.
- *Influence the continuation or discontinuation of a policy, program, or project.* Evaluation findings provide the basis for informed decision making and can support evidence-based programming and policy making.
- *Account to stakeholders for public expenditures, policies, programs, or project results.* Evaluations provide information to taxpayers, policymakers, donors/funders, the private sector, civil society, the media, and other stakeholders about how a program is performing and how well funds are being used to achieve intended results.

- *Help communicate about, and build support for, a policy, program, or project.* If an evaluation yields positive results, it may be used to support promotion or advocacy efforts and serve as the basis for communicating about the benefits or effectiveness of a policy, program, or project, to ensure its continuation.

The intended uses of evaluation findings thus often extend beyond implementation decisions regarding a specific policy, program, or project to more broadly influence decisions at the agency or sector levels. An example of how evaluation information can impact public policy decisions is shown in Box 3.

Box 3: Implications of the International Food Policy Research Institute’s Evaluation of *Progresa* (1998 – 2000)

In 1998, the International Food Policy Research Institute was asked “to assist *Progresa*’s program administration in determining whether the program was functioning in practice as it is intended to by design.” The Institute’s evaluation had significant policy implications. Because it showed that the program had made a substantial “impact on the welfare and human capital of poor rural families,” the new Mexican government of 2000 under Vicente Fox maintained and expanded the coverage of the program (renamed *Oportunidades*) into urban areas.

Source: Skoufias 2005.

In addition to defining the purpose and rationale of the evaluation, it is necessary to assess if such an evaluation is feasible. Planning an evaluation involves trade-off decisions, with the evaluation manager typically having to weigh the costs and feasibility of various designs against the likely benefits of the findings for program administrators, policymakers, the public, or other audiences. Common problems during this planning stage can be avoided by assessing whether a program, policy, or project can be—

- “*Meaningfully evaluated*”—with clearly defined objectives and attainable data,
- “*Whether the evaluation is likely to contribute to improved performance*”—with some agreement among intended users on how the results will be used (Wholey 2004, p. 35).

Box 4 presents a checklist for assessing whether a program can and should be evaluated and how common pitfalls can be avoided. Although it focuses on program

evaluations, the elements outlined can apply to most evaluation types. Some evaluation plans dictate when, where, and how evaluations should occur, and under what budgets and circumstances. However, if an evaluation plan does not exist, then it is useful to consider the checklist for assessing the feasibility of an evaluation.

Box 4: Feasibility Checklist for Planning a Program Evaluation

Program Design Elements

- The program's objectives, outputs, and activities are clear enough to be able to evaluate progress, results, and/or impact.
- There is adequate and reliable program information available (for example, monitoring data) to engage in an evaluation.
- The program's stage of development is appropriate for the envisioned evaluation; for example, an evaluation of outcomes is *not* being proposed for a program still in its planning stage.
- The program had adequate sustainability and stability. Political and financial will exists to sustain the intervention while the evaluation is being conducted, and the program design is not likely to change abruptly during the evaluation period.

Resource and Logistical Considerations

- The evaluation manager and/or commissioning agency staff have sufficient availability for this evaluation during the proposed timeframe.
- Stakeholders who have an interest in the evaluation will have the appropriate availability and opportunities to participate.
- Sufficient funds are available for the evaluation, or additional funds can be raised or leveraged as needed.

Utility

- There is a high probability that the evaluation will be used to improve the project or program.
- The proposed evaluation design is appropriate, given available resources, and intended users have expressed interest in the findings.

Sources: IEG; adapted from CDC 2011a, 2011b; Imas and Rist 2009; IUCN 2004.

Step 3: Prepare a Work Plan to Manage the Evaluation

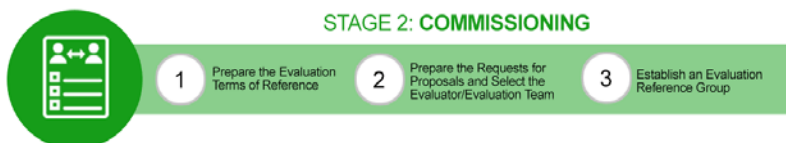
An overall plan for managing the evaluation is an important tool for anyone commissioning an evaluation. Although the more detailed work plan for the evaluation itself might be developed by or in collaboration with in-house evaluators or an external evaluation team of consultants, the manager or evaluation team should identify some key parameters for overseeing the work early on. Considerations during this planning step should focus in particular on clarifying the distribution of roles and responsibilities, the overall timeline, and the budget, as follows (Better Evaluation 2013; CIDA 2004):

- *Roles and responsibilities*—The management plan should spell out how decisions related to the evaluation will be made and describe who is expected to fill key roles.
 - *Decision-making authority*—Who will make which decisions related to the evaluation? What types of structures will be in place to guide decision making? For example, will an advisory group, steering committee, or other formal body be established? If so, will such a group make decisions through voting or consensus or will it simply serve as an advisory body?
 - *Quality assurance*—How will the technical quality of the evaluation be assured—through the reference group or another quality assurance body in the agency?
 - *Evaluation*—Who will conduct the evaluation? Will the evaluation be carried out by relying primarily on in-house capacity or an external consultant or team? Or will a combination of internal staff and external experts work jointly to conduct the evaluation?
 - *Document management*—Who will have responsibility for managing key documents, records, and data? Who will develop the terms of reference (ToR)?
 - *Stakeholder relations*—Who will be in charge of managing stakeholder relations throughout the evaluation?
- *Timeline*—How much time is available to complete the evaluation and report results? Given this timeframe, when should key milestones be completed (that is, developing ToRs, hiring consultants, finalizing evaluation work plan, and so forth)? Extra time should be allotted if possible in between deadlines to allow for unexpected delays and to ensure that the evaluation is delivered according to expectation.

- *Budget*—What resources are available for the evaluation in terms of time, money, and expertise? Is the funding allocation for evaluation built into project or program budgets? Are there other potential resources available? Is the evaluation cost-effective?

Clarifying expectations for all these key areas related to managing an evaluation—in combination with specifying the purpose(s)—will allow the commissioner and/or team to identify the most appropriate evaluation design(s). The process of continuing to refine the evaluation work plan can then be done in collaboration with the evaluator(s). In addition, the evaluation commissioner or team should also consider what information will be needed and how difficult it will be to collect. In some cases, obtaining data and information can be a timely and costly endeavor, and this should be taken into account.

Stage 2: Commissioning the Evaluation



After completing the planning steps to define the purpose and key parameters of the evaluation, the evaluation commissioner or manager will begin the process of commissioning the evaluation. This stage has three key steps:

- Step 1: Prepare the evaluation ToR.
- Step 2: Prepare the requests for proposals and select the evaluator(s)/evaluation team.
- Step 3: Establish an evaluation reference group.

Step 1: Prepare the Evaluation Terms of Reference

The ToR is the document that details the assignment and spells out the requirements and expectations of those commissioning or managing the evaluation. The ToR serves as the basis for a contractual arrangement between those commissioning the evaluation and the consultant(s) and/or in-house staff carrying out the work. However, the ToR should be brief, with the typical length 5-10 pages (IEG 2011a). Administrative annexes can be added to provide supplementary information that can inform the depth and breadth of the evaluation or address key issues, where appropriate. The functions commonly served by a ToR are listed in Box 5.

Box 5: What Is a ToR?

A ToR presents an overview of the requirements and expectations of the evaluation. It provides an explicit statement of the roles, resources, and responsibilities of the evaluators and the evaluation client, as well as resources available for the evaluation.

ToRs provide clearly detailed parameters for:

1. Why and for whom the evaluation is being done
2. What it intends to accomplish
3. How it will be accomplished
4. Who will be involved in the evaluation
5. When milestones will be reached and when the evaluation will be completed
6. What resources are available to conduct the evaluation

Source: IEG 2011a.

Although some aspects of the ToR may vary depending on the hiring organization, local norms, and the specific tasks, most ToRs share certain common characteristics, such as those outlined below, adapted from IEG's guide, *Writing Terms of Reference for an Evaluation: A How-to Guide* (2011a).

- a. *Background knowledge and rationale for carrying out the evaluation.* The ToR should provide background on the "evaluand"—the subject of the evaluation (typically a policy, program, or project)—along with an explanation of why an evaluation is being commissioned at this point in time. The background should include an overview of the subject matter as well as the social, economic, cultural, political, and environmental context of the policy, project, or program in terms of what it aims to achieve and whom it will serve. It should also specify the type of evaluation that is being commissioned.
- b. *Objectives*—The framing and presentation of the objectives of the evaluation is a brief but important section in the ToR. It builds an understanding of the process and expectations for the desired task(s) by succinctly presenting information about why the evaluation is being conducted, its objectives, and its intended users.
- c. *Consultation of stakeholders*—Key stakeholders in the evaluation should be consulted to hear their opinions and recommendations. Evaluation managers or commissioners should ensure that these views are incorporated appropriately in the ToR.
- d. *Scope, approach, and methodology*—The ToR should specify the scope of the evaluation, the approach, and the evaluation plan and propose a methodology. The scope or parameters of an evaluation usually refer to the depth, breadth, and timeline of the evaluation. In some cases, one of the first outputs of the ToR is an evaluation plan that describes the scope, evaluation design,

approach, and methodology; however, this is not always the case. If the evaluation manager is prescribing the methodology, the ToR will include, for example, specific guidance on the data collection strategy to be used. Depending on the research questions and resources available, an evaluation may be conducted as an impact evaluation using randomized control trials or it may be conducted using quasi-experimental methods. If the evaluation is not an impact evaluation, it is likely to take a mixed-methods (quantitative and qualitative) approach. Appropriate methods must always be discussed and determined based on the research questions, what is being evaluated, and the resources available.

This section of the ToR should also take into consideration any geographic or cultural sensitivity that evaluators need to be mindful of, both as part of evaluation design and in undertaking the evaluation. The expected approach should be outlined to set realistic expectations among all stakeholders. The degree to which the evaluator(s) can propose additional or alternative methods for completing the task(s) should also be specified.

- e. *Qualifications of the evaluators*—Specific competencies, skills, and characteristics of the evaluators or the team should be outlined in the ToR and determined beforehand with evaluation stakeholders. Broadly, evaluator competencies should include:
- Understanding of evaluation roles
 - Ability and experience managing and implementing evaluation projects
 - Knowledge of evaluation methods
 - Good interpersonal, writing, and communication skills
 - Cross-cultural competence and sensitivity
 - Understanding of ethical practices.
- f. *Budget*—The cost of the evaluation should be estimated beforehand and the ToR should identify the resources that will be available for the evaluation. The budget should include consulting fees, cost of data collection and analyses, any travel, communication, and any reporting costs associated with undertaking the evaluation. Special attention should be given to factors that might affect the budget; for example, the evaluation of a regional program might require the translation of data collection instruments and reports into one or more languages.

- g. *Schedule and reporting requirements*—The ToR should identify the beginning and the completion date of the evaluation. The document should also include a timeline for the evaluation with key milestones and deliverables such as initial findings, interim findings, and submission of the draft and final reports.
- h. *Communication formats and requirements*—The format in which deliverables should be presented should be articulated in the ToR, along with preferred modes of communication (email, phone, and so forth).

Step 2: Prepare the Request for Proposals and Select the Evaluator/Evaluation Team

Once all the details of the ToR have been reviewed and approved by key stakeholders, the next step is preparing a request for proposal (RFP) and selecting the evaluator or evaluation team. Requests for proposals should include the format for the proposal, the content, the submission process, and deadline for submission. Specific elements may include:

- Topics to be covered
- Format requirements and length
- Number of proposal copies and how these should be submitted (hard copy, electronic copy)
- Criteria on how the proposal will be judged
- Timeline for judging proposals and responding to bidders.

Each organization may have its own internal guidelines on selection and procurement processes that must be followed. It is a good idea to put in place a selection committee that is able to review the bidders' proposed methodologies and that has both knowledge of the issues to be addressed and awareness of the ethical and procedural requirements of conducting the evaluation.

Bidders are usually given the opportunity to ask questions to clarify the RFPs up to a certain deadline. Some organizations have policies that questions should be in writing and that all bidders should receive the answers to questions asked, not only the bidder asking the question. There may also be a bidder's conference where consultants can meet with the team commissioning the evaluation to ask questions (IEG 2011a). Although in most cases a competitive process will be used, there may be cases where the organization decides to hire evaluator(s) on a sole-source basis, if it is consistent with procurement regulations.

It is important to note that the evaluator or evaluation team should be selected based on agreed upon selection criteria to ensure that the evaluator or team has the skills and competencies necessary to undertake the work. Evaluations require that the evaluator or the evaluation team have a good understanding of the purpose of the evaluation as well as good technical expertise and knowledge of the program area and good evaluation skills.

Step 3: Establish the Reference Group

Evaluation managers should consider establishing a reference or evaluation advisory group to ensure that the evaluation incorporates state-of-the art technical thinking from external experts and represents stakeholder interests. The reference group is different from the steering committee, as a steering committee is typically responsible for managing the evaluation, while a reference group is usually comprised of external actors that provide additional knowledge and technical guidance. Some potential functions and value of having a reference group are shown in Box 6. Reference groups are also referred to as advisory or consultation groups and may be tasked with:

- Reviewing the ToR
- Providing advice on the composition of the evaluation team and work plan
- Reviewing and commenting on draft reports.

Box 6: Why Have an Evaluation Reference Group?

There are several good reasons for having and using a reference group for the evaluation. A reference group may:

- Provide needed scientific, programmatic, and related expertise to supplement that of staff.
- Provide insights from program practices and can contribute to a deeper understanding of “what is really going on.”
- Provides legitimacy to the evaluation by providing an independent perspective.
- Contribute to the use of evaluation findings by providing visibility to the evaluation and helping to disseminate findings

Source: Adapted from CDC 2011a, 2011b.

Such a group can help enhance stakeholder ownership on the evaluation and its results, which can lead to a more engaged process with stakeholders feeling more enfranchised and providing more feedback and comments on evaluation findings and draft reports. The group might consist entirely of members who are external to the organization commissioning the evaluation or it might include a mix of internal and external members.

Evaluation reference groups are established to provide advice on an individual evaluation, a series of evaluations, or the evaluation function within an organization. Therefore, the manager of the evaluation should first confirm whether there is an existing advisory body serving this function. If not, the process for establishing such a reference group should include working with key stakeholders to determine what role(s) the group will serve and identifying who should be invited to serve as members. Developing a ToR for the group is useful for clarifying responsibilities and expectations.

The American Evaluation Association has dedicated a journal issue to evaluation advisory groups with further guidance and areas for consideration (see Velure Roholt and Baizerman 2012).

Stage 3: Undertaking the Evaluation



Managing an evaluation is frequently considered more of an art than a science, as the work can take place in complex environments and unpredictable situations can arise. No one can completely plan for everything that may arise, but the best approach is to imagine and plan for possible contingencies and to ensure that the views of all key stakeholders are considered at the outset.

The process of managing the implementation of an evaluation typically includes three steps.

- Step 1: Supervising the evaluation work plan and implementation
- Step 2: Ensuring evaluation quality
- Step 3: Providing feedback on the report.

Step 1: Supervise the Evaluation Work Plan and Implementation

The process of overseeing the actual implementation of the evaluation includes not only supervising an evaluator or an evaluation team but also organizing and managing documents and other materials for the evaluation and maintaining communication with key stakeholders as appropriate.

The evaluation commissioner or manager and selected evaluator(s) should review the ToR together to ensure that there are no questions, comments, or points that need to be renegotiated. If there are, then the evaluation commissioner or manager should decide if the ToR should be amended or not. In addition, both parties should agree on the remuneration for the work so that a contract can be issued. Covering these considerations and establishing a system for progress reporting at all stages of the evaluation will ensure that if problems arise, they are tackled sooner rather than later. Setting up a schedule for check-ins will help facilitate communication on the achievement of evaluation milestones and any issues that may arise.

Both the evaluation manager and evaluator (or evaluation team) have important responsibilities, which need to be met to ensure a high-quality evaluation. Both parties should share an understanding of the ToR and business processes. Each individual can help strengthen the evaluation by carrying out the relevant activities listed in Box 7.

Box 7: Key Roles and Activities to Ensure a High-Quality Evaluation

The evaluation manager should:

- Ensure that the objectives of the evaluation are clear
- Maintain ownership of the study by ensuring that decision-making responsibility is retained and that decisions are made in a timely manner
- Negotiate expectations with stakeholders
- Monitor the progress of the evaluation and provide relevant and timely feedback and guidance to the evaluator and evaluation team
- Be open to suggestions from evaluators on possible solutions if problems arise
- Discuss and ensure agreement on formal and informal communication protocols from the beginning
- Ensure evaluators have full access to information as early as possible
- Meet with evaluators and stakeholders to discuss draft reports and revisions
- Approve the final report and organize a presentation of evaluation findings for stakeholders (UNEG 2005a, 2005b).

The evaluator (or evaluation team) should be encouraged to:

- Commit to conducting the evaluation within the allotted time frame and budget
- Provide regular progress reports to the evaluation manager and communicating problems that require the attention of the evaluation manager immediately
- Discuss the draft report and correct any errors or misinterpretations
- Respond to comments and finalize the report.

Source: Compiled by IEG.

Before and while the evaluation is being conducted, the manager or commissioner is responsible for organizing the relevant materials and data and ensuring that these

are available to and used by the evaluators as appropriate. The process of identifying and managing key program documents and other materials relevant to the evaluation can be time consuming, so the commissioner or manager should factor in the time and/or support needed to do the following:

- Identify and collect relevant background documents, program materials, organizational charters, memoranda of understanding or other agreements among partners, previous evaluation reports, or other available resources that will help to inform the evaluation
- Compile and prepare relevant existing data such as that related to program activities and participant information
- Brief the evaluator or team on available documents, materials, and data and ensure that these have been provided in an accessible format and that the evaluator(s) understand the context and range of available information
- Monitor the use of key materials and data throughout the evaluation to ensure that these are being drawn on as appropriate, given the study design.

Finally, the commissioner or manager should communicate with key stakeholders regularly. Common activities that comprise this role include:

- Identifying the key individuals and/or stakeholder groups that should be aware that the evaluation is being undertaken or who are expected to have a role in the evaluation (that is, as respondents, peer reviewers, and so forth)
- Drafting and sending general information about the evaluation:
 - A letter at the start describing the evaluation and introducing the team
 - A letter at the end sharing a summary of findings and notifying stakeholders where they can access the evaluation report or learn about follow-up activities
- Coordinating correspondence for targeted respondents; for example, a letter from a high-level official in the commissioning agency might be an effective mechanism for increasing the survey response rate among past program participants.

Step 2: Ensure Evaluation Quality

Quality standards ensure that evaluations provide accurate and useful information and comply with procedural requirements within an organization (for example, communicating findings in a specific way), along with technical requirements and ethical guidelines. Most organizations and agencies will already have established evaluation quality standards, principles, or norms (for example, see Box 8 for the norms for evaluation at the United Nations Development Programme). Sometimes these standards are included as part of the overall evaluation policy (see Stage 1, Step 1); in other cases they are produced as separate documents.

Box 8: United Nations Development Programme Norms for Evaluation

- **Independent**—Management must not impose restrictions on the scope, content, comments, or recommendations of evaluation reports. Evaluators must be free of conflict of interest.
- **Intentional**—The rationale for an evaluation and the decisions to be based on it should be clear from the outset.
- **Transparent**—Meaningful consultation with stakeholders is essential for the credibility and utility of the evaluation.
- **Ethical**—Evaluation should not reflect personal or sectoral interests. Evaluators must have professional integrity, respect the rights of institutions and individuals to provide information in confidence, and be sensitive to the beliefs and customs of local social and cultural environments.
- **Impartial**—Removing bias and maximizing objectivity are critical for the credibility of the evaluation and its contribution to knowledge.
- **Of high quality**—All evaluations should meet minimum quality standards defined by the Evaluation Office
- **Timely**—Evaluations must be designed and completed in a timely fashion so as to ensure the usefulness of the findings and recommendations
- **Used**—Evaluation is a management discipline that seeks to provide information to be used for evidence-based decision making. To enhance the usefulness of the findings and recommendations, key stakeholders should be engaged in various ways in the conduct of the evaluation.

Source: UNDP 2011.

If an organization does not have quality standards already in place, it is useful to consult the standards and principles set forth by the local professional association of evaluators or to refer to the guidelines of OECD's Development Assistance Committee (DAC) as a starting point. These evaluation standards aim to provide a roadmap for ensuring the quality of development evaluations. They are based on the following key evaluation principles: "impartiality, credibility, independence, and usefulness" (OECD 1991). The standards should be used throughout an evaluation to provide quality checks at its different stages (OECD 2010b). In addition, quality evaluation standards should be used (or developed) with input from stakeholders and should be adapted to local and national contexts.

Technical Quality

To ensure the technical quality of an evaluation, evaluation commissioners and managers should ensure that the evaluator (or evaluation team) asks the following questions:

- Are the data collected accurate?
- Are the data relevant to the evaluation questions?
- Is the analysis clear and logical?
- Are there any gaps?
- Are the findings and lessons presented appropriately?
- Are the recommendations cost-effective and actionable?

Ethical Guidelines

Ethical guidelines are sometimes covered by quality standards, or they may be published separately. All evaluations should be conducted in an ethical manner, especially with regard to vulnerable populations. It is key for evaluation managers to understand the local context in which the evaluation will be conducted and to communicate expectations for ethical practices.

In many cases conflicts of interest may arise; for example, individuals who may have worked on a project or program years before may seem like good candidates to be involved in an evaluation. However, managers need to be careful because these types of situations can yield biased results. If the manager is not aware of the person's past involvement, the person being approached for the work should be forthcoming and offer to recuse him or herself from conducting or working on the evaluation.

Evaluation commissioners or managers may have a set of quality standards that their organization already subscribes to. In these cases it is important to ensure

that the evaluator (or evaluation team) subscribes to them, and that they are acceptable to the reference group and other stakeholders.

If neither party has a preferred set of ethical guidelines, the evaluation manager can review those of various evaluation professional associations or other organizations that routinely practice, or guide the practice of, evaluations, such as United Nations Evaluation Group, the American Evaluation Association, the African Evaluation Association, the Canadian Evaluation Society, the European Evaluation Society, the Global Environment Facility, and so on (OECD 2010b).

Step 3: Provide Feedback on the Inception, Draft, and Final Reports

The evaluation commissioner or manager should plan to oversee and provide feedback on all the reporting for the evaluation. Typically, the evaluator(s) will provide an inception report with a detailed evaluation plan near the start of the evaluation period, and then draft and final reports once the evaluation has been completed. When each report has been written, it is critical for the evaluation commissioner or manager, the reference group, and other key stakeholders as appropriate to review it and provide feedback. This step will help ensure that all the intended users of the findings will receive the information they need.

If there are significant issues at any stage of the reporting, then the evaluator and manager should re-assess the process and develop a plan to address them. Often the problems that arise are technical, ethical, or procedural:

- A technical issue, for example, is where a report makes a strong statement about causality, without having tested for causality; it could make statements about trends based on a period too short to generalize from; it could claim a significant result when in fact the result was not statistically significant; or there could be statistical results based on an unrepresentative sample or on analysis that leaves out key variables.
- Ethical issues relate to whether proper evaluation protocols were followed when conducting interviews and other data collection and analysis activities.
- Procedural issues can usually be addressed more easily and may not risk the integrity and validity of the study. These tend to relate to whether the report is clear and logical and whether findings are expressed appropriately and in an appropriate format required by the evaluation manager.

Any of these kinds of issues will require that the evaluation team figure out where mistakes have been made and whether they can be corrected; otherwise, the

evaluation will likely not be useful and could do more harm than good, especially if there were technical and ethical concerns.

The evaluation commissioner or manager is ultimately responsible for the quality of the evaluation and should plan to review the relevance and accuracy of any reports and compliance with the ToR. However, it is also important to note the need to safeguard the independent view of the evaluation team. The team leader will determine the evaluation's conclusions and recommendations and the evaluation manager must decide what actions to take if there is strong disagreement from stakeholders. One option is to add content about the disagreement to the evaluation report and possibly a management response, such as from a program office or from a specific agency or department of a government as relevant (WFP 2013).

Specific responsibilities for reviewing the evaluation are best established during the planning stage of the evaluation. One common approach is that after the evaluation has been reviewed by the evaluation manager, the team shares the report with the reference group and other stakeholders for feedback and comments. As part of this review, all parties should ensure that the recommendations in the report are actionable. In many cases, evaluators may provide recommendations that are not practical in the circumstances and context in which a given project is operating, or the recommendations may simply not be able to be implemented because of constraints such as budget, time, and so forth.

To address this issue, some organizations have put out guidelines on how to write good evaluation recommendations. The Independent Evaluation Group of the World Bank Group created a 12-point checklist on writing good recommendations (see Appendix B).

Once the evaluation manager, the reference group, and other key stakeholders have reviewed, commented on, and discussed feedback internally, a meeting should be scheduled with the evaluator or evaluation team to discuss feedback and comments and to jointly decide on changes that will be made or incorporated to finalize the evaluation and disseminate the findings to a broader audience.

Stage 4: Reporting and Follow-Up



Even after the evaluation report is final, the evaluation commissioner or manager still needs to make critical decisions about how to most effectively communicate and disseminate findings and ensure that the results are used to improve future practice.

- Step 1: Communicating and reporting results.
- Step 2: Following up on use of evaluation recommendations.
- Step 3: Wrapping up the evaluation.

Step 1: Communicate Results

After the reference group and stakeholders have approved the final report, the evaluation commissioner or manager should consider and discuss how the findings of the report will be presented to other stakeholders, particularly if the findings are sensitive or controversial.

Evaluation commissioners or managers should also consider the following when crafting a communication and dissemination strategy and timeline for an evaluation:

- How will the findings of the evaluation be reported and disseminated to relevant audiences (publications, presentations, roundtables, social media, and conferences)?
- How will the findings be shared with various groups of stakeholders with diverging points of view?
- When is the best timing for communicating the findings to ensure their use?
- Should the report, or summaries, be translated to enable all stakeholders to make use of the findings and the recommendations?

Evaluation commissioners or managers should consider what reporting mechanisms would resonate best with stakeholders. For example, if evaluation findings are applicable to an organization or institution, organizing conferences, meetings, or forums may work best. In this situation, the evaluation manager may want to hire a communication specialist to help facilitate the events. This person should have a number of skills, including a degree of cultural competence in communicating with diverse groups, some knowledge of the subject matter if possible, and appropriate media and information technology skills. However, if the audience is much wider and geographically dispersed, then presenting and disseminating findings may entail using video or webinar, publishing synthesis reports, posting materials on the Web, issuing press releases, and possibly organizing conferences.

Communication specialists can also be helpful in cases where the evaluation manager needs to communicate findings to sensitive stakeholders or to mitigate potential adverse reactions to negative feedback. Specialists can advise on how to present findings in a constructive way that encourages engagement with and use of those findings. One way to encourage engagement is to cast the findings in forward-looking ways and to focus on implications for the future.

The evaluation and its findings may also be presented at professional conferences as a way of sharing knowledge and lessons broadly and communicating with experts working in the area the evaluation addressed.

Ultimately, evaluation managers must realize that the communication and dissemination of results requires analyzing stakeholders, knowing the audiences as much as possible, and knowing how they may react to findings in order to devise a strategy that ensures findings reach the right people at the right time and are used.

Step 2: Follow Up on Use of Evaluation Recommendations

Once findings are disseminated, evaluation commissioners and managers need to be responsive to the needs of various audiences and stakeholders. It is important that managers and evaluators be as transparent as possible in answering questions, in particular on the methodology and approach to the evaluation and the implementation process.

To ensure that the findings and recommendations from evaluations are used, managers should be sensitive to a number of factors when disseminating and

following up on the report. Specifically, they should:

- Ensure the timely delivery of the report and responses to questions or comments on the findings; avoid information overload and consider the format for delivery for the various stakeholders.
- Discuss findings and recommendations with the programs/projects affected and, to the extent possible, agreed on these with the program managers/funders/stakeholders before the disclosure.
- Ensure that the recommendations are timely and actionable; frame recommendations in a digestible and actionable manner.
- Actively promote follow-up and use of the evaluation findings by finding champions for the recommendations (through workshops, publications, seminars, and so forth).

In some cases, ensuring that program managers accept and adopt recommendations can take some influencing and negotiation and may not happen quickly. However, it is important to work with program managers to reach an agreement.

In addition, some countries and organizations have instituted systems or mechanisms which allow for following up and tracking the adoption of recommendations. Appendix C presents two examples of such systems: the World Bank's Independent Evaluation Group's system—the Management Action Record—and Mexico's Follow-Up Mechanism for the Use of Evaluations.

Step 3: Wrap Up the Evaluation

Although the implementation of recommendations and follow-up activities might continue over a longer term, the evaluation commissioner or manager will need to formally end the evaluation task itself. The completion or closing of an evaluation should include clear decisions about the storage and release of any reports and data over the longer term. If the manager has not already established who will have access to reports and how data and other materials will be archived, then these activities need to be undertaken as part of a wrap-up step.

One key decision is who will have access to the evaluation findings in the future. Many organizations have policies in place to determine the level of public access to evaluation information. For cases where such guidelines have not yet been established, local professional associations for evaluators can serve as useful resources to develop such a policy. The American Evaluation Association (AEA)

Program Evaluation Standards promote the public disclosure of findings given the “evaluator’s obligation to serve the broader public who benefit from both the program and its accurate evaluation,” but the AEA acknowledges that its standards will not be appropriate in all countries and contexts (AEA 2004; Imas and Rist 2009, p. 507).

The evaluation commissioner or manager should consult the standard policies and procedures of the organization and determine whether and how one or more reports from the evaluation will be accessible to the public or designated audiences. In addition, similar guidance is needed regarding the archiving and management of data or other artifacts from the evaluation.

Typically, the commissioning agency will maintain long-term control over the data and artifacts from an evaluation rather than the evaluation consultant or team. Specifying roles, responsibilities and ownership throughout all stages of the evaluation will help to facilitate the closure of the evaluation and help to ensure the long-term value of the findings and data.

Decisions and protocols for the evaluation data should cover at least the following considerations (BetterEvaluation 2013; FAO 1998):

- *Secure data storage*—What processes are needed to protect electronic and hard copy data in all forms, so that these are not accessed without authority or damaged? These data could include questionnaires or other hard-copy materials, interview tapes, and electronic files.
- *Backing up data*—Are there protocols or automated processes in place to back up data so that the risk that they will be lost or corrupted is reduced? One approach is to store a duplicate set of data offsite.
- *Archiving for future use*—Has a data set been specifically cleaned and stored for future use, to support further analysis and research? Sometimes a cleaned data set is archived for broader research use in which sensitive identifying information has been removed.
- *Data access and dissemination*—Has the authority for managing and/or releasing the data in the future been clearly specified? What process should be used for someone seeking to use the data to verify the evaluation’s findings or to use the data as a basis for further research? If the organization does not already have clear guidelines in place, the procedures for control of the data and permissions for access should be defined as clearly as possible.

Appendix A: Sample National Evaluation Policies

United Kingdom's Official Policy:

Some countries have formal, legally mandated and written evaluation policies, and others have informal practices. Aside from South Africa, another example of an evaluation policy is that of the United Kingdom's Department for International Development (DFID). This policy articulates the organization's commitment to learning from evaluation, measuring results, and accountability. It also sets out guiding principles, criteria, and standards, as well as its evaluation approaches (see Box A.1).

Box A.1: U.K. Department for International Development's Policy for the Evaluation of Official Development Assistance

Purpose: Evaluation is a significant part of DFID's strategy to using evidence more effectively to achieve maximum impact from its development assistance. According to DFID, "Evaluation has a key role in generating evidence and learning about what is working in development and what is not; it can identify better ways of doing things, allow for course corrections of programs to improve effectiveness; ensure that lessons are learned during the development process and resources shifted to where they are most effective; and improve the ability to respond to change."

Definition of Evaluation: DFID subscribes to the Organisation for Economic Co-operation and Development–Development Assistance Committee internationally agreed definition of evaluation, "a systematic and objective assessment of an ongoing or completed project, program or policy, its design, implementation and results in relation to specified evaluation criteria."

The core guiding principles for DFID's policy are:

- **Independence:** When possible, evaluations are to be carried out by specialists who are independent of those designing and implementing the intervention.
- **Transparency:** Lessons are to be shared publicly, after allowing for legal obligations and protection of confidential information.
- **Quality:** Evaluation strategies and plans should be clear and actionable and have a budget, with approaches and methods reflecting the best available in the profession drawing on latest research, drawing on the latest research.

- **Utility:** Evaluation process and design has to keep “use” in mind – ownership by users in the process.

Evaluation Approaches: The choice of the actual evaluation approach in DFID is guided by four factors: the content within which the initiative is operating, the attributes of the program, the evaluation questions posed and the criteria being used to judge these components. There are two broad evaluation approaches:

- **Theory-based approach:** This focuses on understanding the characteristics of development interventions—what it is about a particular approach that works. It is suitable for complex initiative where there are multiple assumptions and lines of inquiry, and it typically requires more qualitative than quantitative data.
- **Impact approaches:** These focus on establishing cause and effect. This approach is better suited for single lines of inquiry and is most useful for pilot initiatives and phased roll-out. Where one wants to see where the initiative is working or not.

Source: U.K., Department for International Development 2013.

Chile’s Informal Policy:

Chile is an example of a country with an informal evaluation policy. Although Chile does not have an evaluation policy framework, it carries out evaluations under its system of management control and results-based budgeting. The government originally launched the evaluation program after a 1996 political agreement with Congress, aiming to enrich the information available for budgetary decision making. The creation of the evaluation program was in part a response to demands from the legislature to have better quality information about government programs, to increase its ability to affect budget decisions. In the beginning, the evaluations were exclusively rapid desk evaluations, which were primarily based on secondary sources. Later, the more profound impact evaluations were added. In 2003, several laws were passed with the aim of strengthening the evaluation program (Box A.2).

Box A.2: Chile's Informal Evaluation Policy

Chile's government evaluation program is one of the six instruments of its monitoring and evaluation systems, known as the System of Management Control and Results-based Budgeting of the Budget Direction Office (Dirección de Presupuesto). Other instruments include:

1. Strategic definitions (Definiciones Estratégicas)
2. Management Improvement Program (Programa de Mejoramiento de Gestión)
3. Performance Indicators (Indicadores de Desempeño)
4. Comprehensive Management Reports (Balance de Gestión Integral)
5. Competitive Fund (Fondo Concursable).

Objectives of the Evaluation Program: The annual budget law provides the legislative framework for the evaluations. The objectives of the program are the improvement of the efficiency of resource allocation and of policies, program management, and accountability. Evaluations have been institutionalized through an annual budget law and their routine use in budgeting.

Chile carries out three main types of evaluations and reviews:

- Government program evaluations (Evaluación de Programas Gubernamentales)
- Impact evaluations (Evaluación de Impacto)
- Comprehensive spending reviews (Evaluación Comprensiva del Gasto).

Who Is Involved: At first these evaluations were done by panels composed of both public servants and independent evaluators. Since 1998, the panels are composed only of external evaluators. They are chosen in a public tender to ensure transparency, objectivity, and independence of the evaluations. After 2001, DIPRES incorporated the Impact Evaluations, to evaluate the actual impacts caused by the program on the situation the government is attempting to correct or improve.

Source: World Bank 2005.

Appendix B: A Guide to Developing Good Recommendations in Evaluation Reports

Recommendations should be:

- Relevant to the object and purposes of the evaluation.
- Based on the evidence and conclusions, logically related to the key findings and appropriate for what has been learned.

Developed in an iterative manner in consultation with management and operational staff, always understanding that the final recommendation represents the views of the evaluation team and cannot be dictated by management.

- Clearly stated in simple, straightforward language and written in a prescriptive manner to facilitate implementation (neither so broad that they become motherhood or so detailed that they unnecessarily take degrees of freedom away from management).
- Prioritized in terms of urgency and timing.
- Clearly identifying the unit that is supposed to take action on the recommendation.
- Actionable, feasible, and reflecting an understanding of potential constraints to implementation.
- Capable of being monitored using verifiable indicators, and time bound—when specific, monitorable actions are expected to be taken.
- Relatively few in number, generally five to six per evaluation (with not too many subrecommendations), covering resource allocation, financing, planning, implementation, and monitoring and evaluation.
- Accounting for the least costly alternative to achieve the objectives and for whether the benefits resulting from the recommendations exceed the likely cost of implementation, based on existing evidence or the literature.

Sources: Among other documents, this checklist draws on material from ADB 2008, 2009; UNEG 2010, p. 6; UNEG 2005, Standard 4.16; Hendricks and Handley 1990; USAID 2010; CIDA 2004; Perrin 2009.

Appendix C: Examples of Follow-Up Mechanisms for Tracking the Adoption of Recommendations

Some organizations, such as the World Bank Group’s Independent Evaluation Group (IEG), have instituted a system that allows for follow-up on the adoption of recommendations. IEG is mandated to report periodically to the World Bank Group’s Board on the adoption of its recommendations in efforts to promote accountability and learning within the Bank Group.

IEG carried out a series of case studies on some evaluations to determine the factors that contributed to the adoption of recommendations. The case studies found that the following factors matter:

- In-depth discussion with management during the drafting of recommendations
- Credibility of the evaluation results
- Sense of shared ownership over the evaluation and the findings
- Quality of the recommendations, specifically in terms of their timeliness, how actionable they are, and their cost-effectiveness, clarity, and coherence
- Advocates/champions supporting the adoption of the recommendations
- Institutional incentives and accountability for adopting the recommendations.

Based on the findings from the case studies, IEG proposed a series of reforms to its process for drafting and following up on recommendations. Figure C.1 shows the process and the reform recommendations

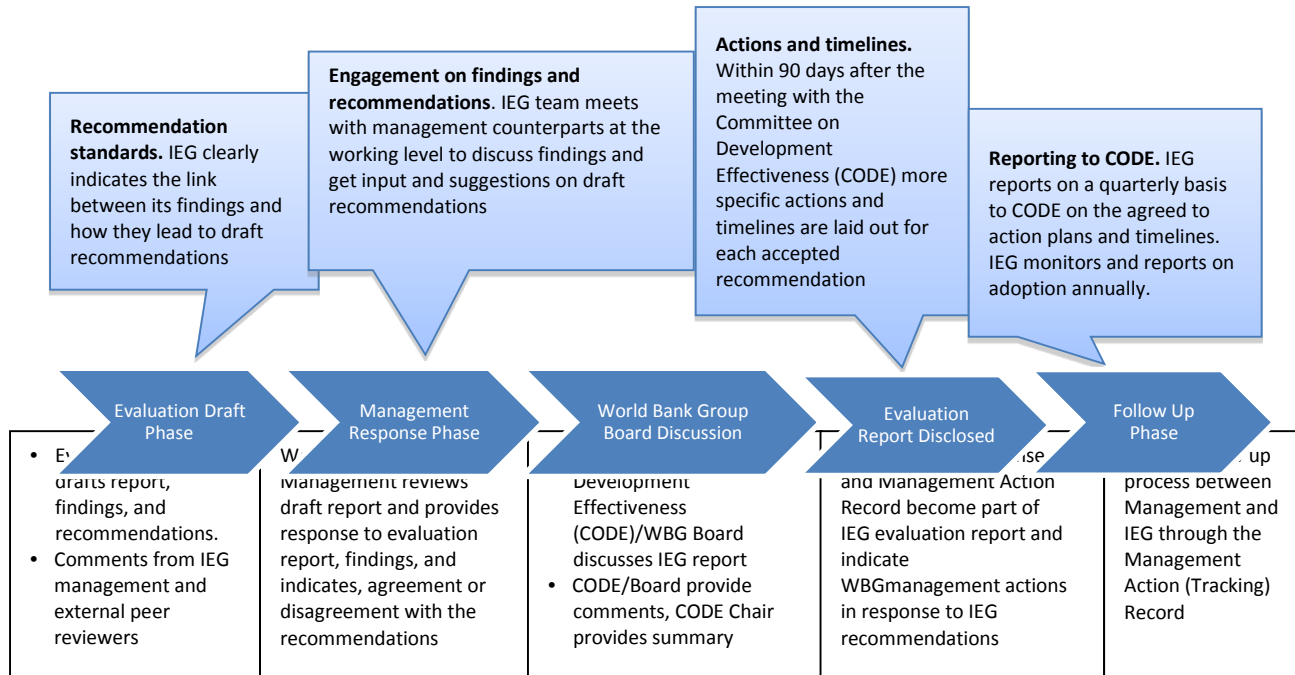
Other agencies and organizations have devised other ways of following up on evaluation recommendations that fit their needs, as depicted by the case of Mexico (Figure C.2).

IEG’s 2012 annual report indicated that the reformed process is yielding “more constructive engagement with Bank Group management and a sharper focus on IEG recommendations” (IEG 2012). The report also noted that a piloting of the process on three evaluation reports has yielded positive results, specifically

creating clearer links between the evaluation findings and the proposed recommendations, which has encouraged greater adoption of recommendations and more ownership (IEG 2012).

An example of another follow-up system is Mexico's Follow-Up Mechanism for the Use of Evaluations. The "Mechanism," as it will be referred to in this text, was devised by the Mexican federal government's National Council for the Evaluation of Social Development Policy (*Consejo Nacional de Evaluación de la Política de Desarrollo Social* - CONEVAL), in conjunction with the Secretary of Finance and Budget (*Secretaría de Hacienda y Crédito Público* – SHCP) and the Secretary of Public Administration (*Secretaría de la Función Pública* - SFP). Its purpose is ensure the use of findings and recommendations of external evaluations to improve Mexico's social programs (UNDP 2013).

Figure C.1: IEG’s Recommendation and Follow-Up Process



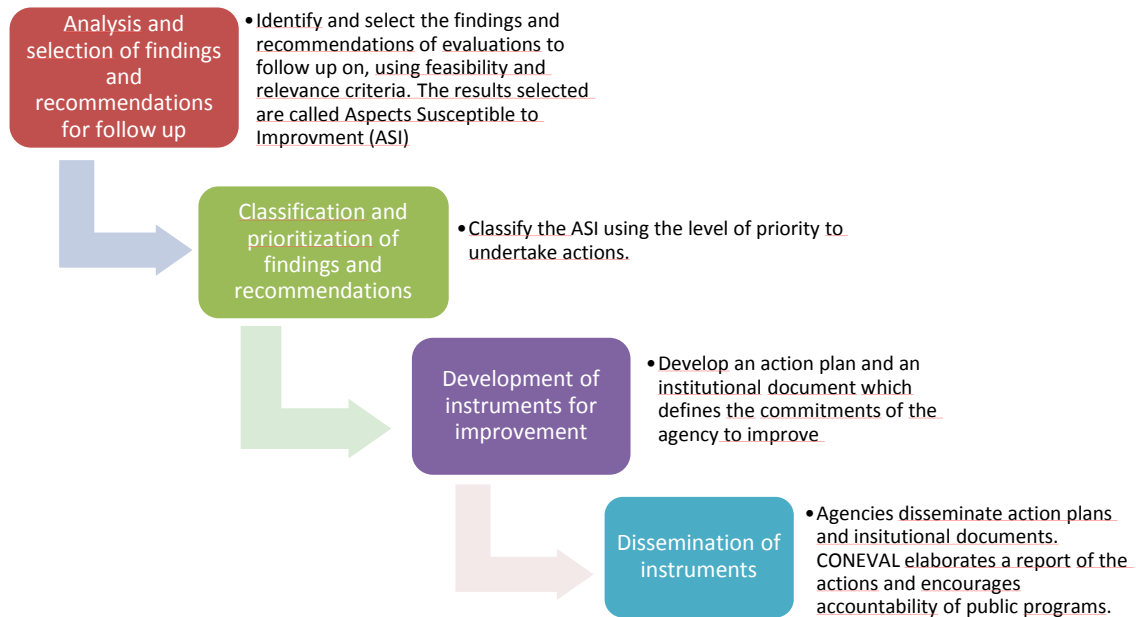
The purpose of the Mechanism is twofold:

- To establish a general procedure to monitor areas identified through the evaluations where there are agreed upon actions for improvement
- Disseminate the commitments made by program operators to improve the public programs evaluated

As observed in Figure C.2, the mechanism has four stages:

- **Stage 1—Select Findings and Recommendations**—Evaluators and program operators select the findings and recommendations they will address by using feasibility and relevance criteria. These are then called Aspects Susceptible for Improvement (ASI)
- **Stage 2—Classify and Prioritize**—Evaluators and program operators define the responsible areas where actions must be taken to improve the program. Other actors from other parts of the program are involved at this stage, as improving a program typically involves the interaction of several areas (operation, budgeting, and so forth) In addition, ASIs are classified as being high, medium, or low priority.
- **Stage 3—Develop an Action Plan**—Development of instruments and elaboration of an action plan by the agency making the improvements and CONEVAL. The action plan documents and defines the commitments of the agency for the improvements of the program.
- **Stage 4—Disseminate the Plan**—Dissemination and publishing of the action plans on the web pages of CONEVAL and the agency/organization involved. This is meant to encourage the agency to carry out the agreed upon actions to improve the program. The agency/organization also agrees to report its advances on each action it has committed to undertake at specific dates.

Figure C.2. Follow-Up Mechanism for the Use of Evaluations



Source: de la Garza and Niembro 2013.

Mexico's Mechanism encourages the participation of several actors, including those in program units, evaluation units, and others tasked with other important aspects of the programs evaluated, such as budgeting, communications, and so forth. This follow-up process requires continual communication as well between evaluators and program operators (UNDP 2013).

The Mechanism has motivated agencies to take concrete actions to improve programs. In Mexico, the evidence suggests that evaluations are becoming a key factor in the improvements of policies and programs. CONEVAL has measured a number of improvements. Figure C.2 categorizes the actions taken and the changes implemented in 2011 as a result of follow-up on evaluation findings and recommendations.

Table C.1: Actions Implemented to Improve Social Development Programs

Activity	2011-2012 Actions Participation (%)	
Improved activities or program processes	163	71
Modify program's services or goods	27	12
Substantially reorient the program	35	15
Increase or reallocate the program	3	2
TOTAL	228	100

Source: CONEVAL. Informe de Seguimiento a los Aspectos Susceptibles de Mejora de los Programas y Acciones Federales de Desarrollo Social 2011-2012.

Table C.1 shows that the Mexican follow-up mechanism has achieved an increase in the systematic use of evaluations and their findings.

Bibliography

ADB (Asian Development Bank). 2009. *Annual Report*. Manila: ADB.

———. 2008. *Annual Report*. Manila: ADB.

AEA (American Evaluation Association). 2004. "American Evaluation Association Guiding Principles for Evaluators." <http://www.eval.org/p/cm/ld/fid=51>

Bell, B. James. 2004. "Managing Evaluation Projects." In *Handbook of Practical Program Evaluation*, eds. Wholey, Joseph S., Hatry, Harry P. and Newcomer, 2nd. Ed. San Francisco: Jossey-Bass.

BetterEvaluation 2013. "Manage an Evaluation or Evaluation System." http://betterevaluation.org/plan/manage_evaluation

CIDA (Canadian International Development Agency). 2004. "How to Perform Evaluations." [http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUIImages/Performance%20Review%20Branch%202/\\$file/Model_Evaluation_Workplans.pdf](http://www.acdi-cida.gc.ca/INET/IMAGES.NSF/vLUIImages/Performance%20Review%20Branch%202/$file/Model_Evaluation_Workplans.pdf)

Catholic Relief Services. 2011. "Monitoring & Evaluation." Chapter 10 in *Institutional Strengthening: Building Strong Management Processes*.

CDC (Centers for Disease Control and Prevention). 2011a. *Developing and Using an Evaluation Consultation Group*. Atlanta, GA: CDC Division of Nutrition, Physical Activity, and Obesity.

———. 2011b. *Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide*. Atlanta, GA: CDC.

Consejo Nacional de Evaluación de la Política de Desarrollo Social. 2013. *Informe de Seguimiento a los Aspectos Susceptibles de Mejora de los Programas y Acciones Federales de Desarrollo Social 2012-2013*. México, D.F. CONEVAL.

———. 2012. *Informe de Seguimiento a los Aspectos Susceptibles de Mejora de los Programas y Acciones Federales de Desarrollo Social 2011-2012*. México, D.F. CONEVAL.

DFID (U.K. Department for International Development). 2013. "International Development Evaluation Policy."
<https://www.gov.uk/government/publications/dfid-evaluation-policy-2013>

[de la Garza, Thania, and Rasec Niembro. 2013. *Incentives in the Use of Evaluations as a Tool for Improving Public Policy*. In United Nations Development Program. *Solutions Related to Challenges of Independence, Credibility and Use of Evaluation, Proceedings from the Third International Conference on National Evaluation Capacities*. 30 September – 2 October 2013, São Paulo, Brazil.](#)

European Commission. 2008. Internal Charter for the Evaluation Function in DG FCIN. Directorate General Economic and Financial Affairs, Resources, Strategic management and control, inter-institutional relations. European Commission: Brussels.
http://ec.europa.eu/economy_finance/publications/publication13113_en.pdf

FAO (United Nations Food and Agriculture Organization). 1998. *Guidelines for the Routine Collection of Capture Fishery Data*. FAO Fisheries Technical Paper. Prepared at the DANIDA/FAO Expert Consultation.
<ftp://ftp.fao.org/docrep/fao/003/x2465e/x2465e00.pdf>

Hawkins, Penny. 2012. "International Program for Development Evaluation: Managing Evaluations." Workshop, Ottawa.
<http://www.oecd.org/dac/evaluation/dcdndep/35857765.pdf>

[Hendricks, Michael, and Elisabeth A. Handley. 1990. "Improving the Recommendations from Evaluation Studies." *Evaluation and Program Planning* 13\(2\): 109-17.](#)

Imas, Linda G. Morra, and Ray Rist. 2009. *The Road to Results: Designing and Conducting Development Evaluations*. Washington DC: World Bank.

IEG (Independent Evaluation Group). 2012 *Annual Report*. Washington, DC: World Bank.

———. 2011a. *Writing Terms of Reference for an Evaluation: A How-To-Guide*. Washington, DC: World Bank.

———. 2011b. *Annual Report. Results and Performance of the World Bank Group*. Washington, DC: World Bank.

IUCN (International Union for the Conservation of Nature and Natural Resources). 2004. "Managing Evaluations: A Guide for IUCN Programme and Project Managers." Gland, Switzerland: IUCN M&E Initiative.
http://cmsdata.iucn.org/downloads/handbook_eng.pdf

Ivins, Ingrid, and Helena Hwang. 2013. "Building Monitoring and Evaluation Capacity in the Republic of Yemen." PREM Notes: Special Series on the Nuts and Bolts of M&E Systems No. 24, World Bank, Washington, DC.

Kusek, Jody Zall and Ray Rist. 2004. *Ten Steps to a Results-Based Monitoring and Evaluation System: A Handbook for Development Practitioners*. Washington DC: World Bank. <http://preval.org/files/00804.pdf>

New Zealand Ministry of Foreign Affairs and Trade Aid Programme. 2011. "Activity Evaluation Operational Policy." New Zealand: New Zealand Aid Programme.

OECD-DAC (Organisation for Economic Co-operation and Development - Development Assistance Committee). 2010a. "Glossary of Key Terms in Evaluation and Results Based Management." Paris: OECD.
<http://www.oecd.org/development/peer-reviews/2754804.pdf>

———. 2010b. "Quality Standards for Development Evaluations." DAC Guidelines and References Series. Paris: OECD.
<http://www.oecd.org/development/evaluation/dcdndep/36596604.pdf>

———. 2008. "DAC Guiding Principles for Aid Effectiveness, Gender Equality and Women's Empowerment." Paris: OECD.

———. 2006a. "Evaluation Systems and Use: Working Tool for Peer Reviews and Assessments." Paris: OECD.
<http://www.oecd.org/dac/evaluation/dcdndep/35857765.pdf>

———. 2006b. "Guidance for Managing Joint Evaluations."
<http://www.oecd.org/dac/evaluation/dcdndep/37484787.pdf>

———. 2001. "Evaluation Feedback for Effective Learning and Accountability." Paris: OECD. <http://www.oecd.org/development/evaluation/2667326.pdf>

———. 1992. "Development Assistance Manual: DAC Principles for Effective Aid." Paris: OECD. <http://www.oecd.org/dac/evaluation/dcdndep/35019650.pdf>

———. 1991. "DAC Principles for the Evaluation of Development Assistance." Paris: OECD. <http://www.oecd.org/development/evaluation/2755284.pdf>

Perrin, Burt. 2009. The Independent Advisory Committee on Development Impact. Review of the Quality of DFID Evaluations: A Delicate Balancing Act.

Republic of South Africa, Department for Performance Monitoring and Evaluation. 2011. *National Evaluation Policy Framework*. http://www.thepresidency.gov.za/MediaLib/Downloads/Home/Ministries/National_Evaluation_Policy_Framework.pdf

Segone, M. (ed.). 2008. *Bridging the Gap: the Role of Monitoring and Evaluation in Evidence-Based Policy Making*. New York: UNICEF.

Skoufias, Emmanuel. 2005. "Progesa and Its Impacts on the Welfare of Rural Households in Mexico." International Food Policy Research Institute, Research Report 139, Washington DC. <http://www.ifpri.org/publication/progesa-and-its-impacts-welfare-rural-households-mexico>

Soares, F., R. Ribas, and R. Osorio. 2007. "Evaluating the Impact of Brazil's Bolsa Familia: Cash Transfer Programs in Comparative Perspective." *Latin American Research Review* 42(2): 173-190. https://muse.jhu.edu/login?auth=0&type=summary&url=/journals/latin_american_research_review/v045/45.2.soares.html

Sonko, R., A. Berhanu, and R. Shamu. 2011. "Key Considerations for Managing Evaluations." Washington DC: PACT South Africa. Retrieved from Better Evaluation.org: <http://betterevaluation.org/sites/default/files/Key%20Considerations%20for%20Managing%20Evaluations.pdf>

UNDP (United Nations Development Programme). 2011. "The Evaluation Policy of UNDP." Executive Board Document. <http://web.undp.org/evaluation/>

UNEG (United Nations Evaluation Group). 2005a. "Norms for Evaluation in the UN System." New York: United Nations.

UNEG (United Nations Evaluation Group). 2005b. "Standards for Evaluation in the UN System." New York: United Nations.

http://www.iom.int/jahia/webdav/site/myjahiasite/shared/shared/mainsite/about_iom/eva_techref/UNEG_Standards_for_Evaluation.pdf

United States Agency for International Development, Catholic Relief Services, and American Red Cross. 2008. *Monitoring and Evaluation ShortCuts: Managing and Implementing and Evaluation*.

VeLure Roholt, R., and M. Baizerman (eds.) 2012. Evaluation Advisory Groups. *New Directions for Evaluation* 136.

WFP (United Nations World Food Programme). 2013. "Monitoring and Evaluation Guidelines." WFP Office of Evaluation, Rome.
<http://www.wfp.org/content/monitoring-and-evaluation-guidelines>

Wholey, Joseph S. 2004. "Assessing the Feasibility and Likely Usefulness of Evaluation." In *Handbook of Practical Program Evaluation*, 2nd ed. Wholey, Joseph S., Harry P. Hatry, and Kathryn E. Newcomer (eds). San Francisco: Jossey-Bass.



THE WORLD BANK GROUP

1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.
Telephone: 202-477-1234
Facsimile: 202-477-6391
Internet: www.worldbank.org

Independent Evaluation Group
Communications and Strategy (IEGCS)
E-mail: ieg@worldbank.org
Telephone: 202-458-4497
Facsimile: 202-522-3125

