



# EVALUATION GUIDELINES FOR FOREIGN ASSISTANCE

Planning and Performance Management Unit  
Office of the Director of U.S. Foreign Assistance  
Final Version: March 25, 2009

## TABLE OF CONTENTS

INTRODUCTION.....	1
<b>PART 1: EVALUATION CONCEPTS, ISSUES AND INSTITUTIONAL RESPONSIBILITIES .....</b>	<b>3</b>
<b>1.1 Meaning and Core Issues for Evaluation .....</b>	<b>3</b>
1.1.1 Definitions of Monitoring and Evaluation .....	3
1.1.2 Core Issues for Evaluation.....	4
<b>1.2 Objectives of Foreign Assistance Evaluations .....</b>	<b>6</b>
1.2.1 Improved Performance .....	6
1.2.2 Accountability .....	6
1.2.3 Informed Decision-Making.....	6
1.2.4 Lessons from Experience .....	7
<b>1.3 Evaluation Principles for Foreign Assistance .....</b>	<b>7</b>
1.3.1 Usefulness .....	7
1.3.2 Credibility .....	8
1.3.3 Independence.....	8
1.3.4 Capacity Building.....	9
<b>1.4 Foreign Assistance Evaluation Categories.....</b>	<b>9</b>
1.4.1 Project Performance Evaluations .....	9
1.4.2 Project Impact Evaluations .....	10
1.4.3 Program Evaluations.....	11
1.4.4 Global/Regional Impact Evaluations .....	12
1.4.5 USG Country Assistance Evaluations.....	12
1.4.6 Organizational Evaluations .....	13
1.4.7 Special Evaluation Studies.....	13
<b>1.5 Evaluation Modalities.....</b>	<b>14</b>
1.5.1 Internal Evaluations .....	14
1.5.2 External (Independent) Evaluations .....	14
1.5.3 Participatory Evaluations.....	15
1.5.4 Collaborative (Joint) Evaluations .....	15
<b>1.6. Institutional Responsibilities for Evaluation.....</b>	<b>16</b>
1.6.1 Post/Mission .....	16
1.6.2 Bureaus.....	17
1.6.3 Central Evaluation Units.....	17
1.6.4 Office of the Director of U.S. Foreign Assistance .....	18

<b>PART 2: MANAGING THE EVALUATION PROCESS .....</b>	<b>20</b>
<b>2.1 Planning Evaluations .....</b>	<b>20</b>
2.1.1 Specifying Objectives .....	20
2.1.2 Formulating Study Questions .....	21
2.1.3 Evaluating Design and Methodology .....	21
2.1.4 Assessing Evaluability .....	22
<b>2.2. Contracting and Finalizing Evaluation Design.....</b>	<b>23</b>
2.2.1 Preparing the Scope of Work.....	23
2.2.2 Criteria for Selecting Evaluators.....	23
2.2.3 Revisiting the Scope of Work and Evaluation Proposal.....	25
2.2.4 Finalizing Evaluation Design .....	25
<b>2.3. Managing Data Collection and Analysis.....</b>	<b>26</b>
2.3.1 Data Collection .....	26
2.3.2 Data Analysis.....	28
2.3.3 Reporting Evaluation Results.....	28
<b>2.4. Review and Follow-up .....</b>	<b>29</b>
2.4.1 Review of the Draft Report.....	29
2.4.2 Forms of Dissemination .....	30
2.4.3 Follow-up.....	31
 <b>ANNEXES</b>	
Annex 1: Legal and Regulatory Mandate for Evaluation.....	A1-1
Annex 2: Evaluation Resources .....	A2-1

## **INTRODUCTION**

These guidelines, jointly developed by the State Department's Office of the Director of Foreign Assistance and USAID, are a step in evolving an overarching framework for evaluating foreign assistance that is shared by State and USAID. The intent of these guidelines is limited. They do not in any way substitute for the formal policies and operational procedures which the State Department and USAID have established for evaluation. Nor do they preclude USAID or State from revising or reformulating their existing guidance in response to their emerging needs and future requirements.

These guidelines are intended for use by State Department and USAID staff responsible for initiating and managing evaluations of foreign assistance programs and projects. These staffers are described as "evaluation managers" in the following pages; they have responsibility for contracting and supervising evaluations but not for conducting them. It is the responsibility of evaluation professionals, and not managers, to design and carry out evaluations using suitable methodologies and knowledge of the substantive field.

The guidelines are divided into two parts:

The first part explains the meaning and core issues of evaluation, identifies objectives of foreign assistance evaluations, and discusses different types of evaluations that can be undertaken. It also discusses evaluation principles that should inform foreign assistance evaluations and the various modalities for evaluation. Finally, it identifies institutional responsibilities for evaluations.

The second part focuses on managing the evaluation process. It describes the steps involved in planning, contracting, managing data collection and analysis, reviewing the draft report and follow-up.

These guidelines are neither comprehensive nor do they cover the technical issues and processes involved in evaluation. Each evaluation should be treated as a distinct research and analytical endeavor, and a standardized approach is neither possible nor desirable. In no way should these guidelines be construed as a manual for conducting evaluations.

Finally, the guidelines contain short annexes on legal requirements and on resources for evaluating foreign assistance programs.

In preparing these guidelines, DFA has reviewed a wide range of evaluation guidance issued by a host of bilateral and multilateral agencies. It has also examined relevant documents published by U. S. Government Accountability Office, American Evaluation

Association, International Development Evaluation Association, the International Organization for the Cooperation in Evaluation, and the Evaluation Network of the OECD's Development Assistance Committee. In addition, it has freely borrowed from USAID's numerous publications including reports on evaluation and evaluation methodology, TIPS and the programming guidance under ADS200 series.

## **PART 1: EVALUATION CONCEPTS, OBJECTIVES AND INSTITUTIONAL RESPONSIBILITIES**

### **1.1 MEANING AND CORE ISSUES FOR EVALUATION**

#### ***1.1.1 Definitions of Monitoring and Evaluation***

Evaluation is defined as a systematic and objective assessment of an ongoing or completed project, program, strategy or policy. It is designed to determine the relevance, effectiveness, efficiency, sustainability and/or impacts of an intervention, strategy or policy. The purpose of an evaluation is to generate credible and useful information that contributes to improved performance, accountability and/or learning from the experience and to assess their effects/impacts and inform decisions about future programming.

In common parlance, monitoring and evaluation are often mentioned together as signified by the acronym M&E. While the two complement each other, they are conceptually and operationally different. Monitoring is a continual process designed to provide information to decision-makers about the progress (or lack of it) of a project, program or policy. It involves checking the progress against pre-determined objectives and targets, and tells us what is happening or has happened. Evaluation, on the other hand, is much more comprehensive. It goes beyond tracking of progress and seeks to identify the underlying factors and forces that affect the implementation process, efficiency, sustainability and effectiveness of the intervention and its outcomes. Table 1 clarifies the distinction between monitoring and evaluation.

Foreign assistance evaluations complement the performance data provided through the “Foreign Assistance Coordination and Tracking System” (FACTS), established by the Office of the Director of U.S. Foreign Assistance (F). This system integrates all USAID and State Department planning and reporting on foreign assistance into a single database. It contains F’s standard (mostly output level) and custom (mostly outcome level) indicators for each operating unit receiving foreign assistance funds, and requires that targets be set for expected results against these indicators. It also captures information on the target populations reached, key partners and other donor and host country activities. FACTS is able to provide information about the performance of U.S. funded programs at global and country levels. As FACTS progresses, its data base will provide a solid empirical foundation to support evaluations of foreign assistance programs.

**Table 1: Monitoring and Evaluation**

MONITORING	EVALUATION
Continuous/periodic	Episodic/ ad hoc
Tracks progress against pre-determined indicators	Validity and relevance of indicators open to review
Focus mostly on inputs and outputs	Deals mostly with outcomes/impacts and with the issues of effectiveness, efficiency, relevance, impact and sustainability
Quantitative methods	Quantitative and qualitative methods
Focuses on what and where	Also focuses on why and how
Usually an internal management exercise	Often multiple stakeholders
Purpose is to improve performance	Objectives are to promote accountability, informed decision-making and learning, as well as to improve performance

**1.1.2 Core Issues for Evaluation**

Broadly speaking, foreign assistance evaluations focus on one or more of the following five sets of issues:

*1. Effectiveness*

Effectiveness refers to the extent to which an intervention’s objectives are achieved or are likely to be achieved. Evaluations seek to determine whether interventions’ services and products are reaching the targeted populations; whether the intended beneficiaries are using them; whether the coverage of the target population is as planned in the project or program design; and whether the intervention is likely to achieve its targets. For example, an evaluation of the effectiveness of a rural credit project for small holders is likely to focus on the issues surrounding the delivery of credit, such as rates of disbursement and recovery, the use of loans by farmers for purchasing the desired agricultural inputs, the likelihood of the project achieving its targets and the proportion of women farmers receiving loans.

*2. Efficiency*

Efficiency measures the output of an intervention in relation to its costs. The most widely used method is to undertake cost- benefit analysis to determine the net benefits of an intervention. Such analyses are most appropriate when (a) reliable data are available and (b) realistic assumptions about benefit streams can be made. Cost-benefit analyses are often problematic in evaluating social or democracy programs, because they require economic quantification of the benefits. For example, what is the economic benefit of saving a human life? How can an economic value be assigned to human rights? Although economists assign economic values to some social variables, the

process remains questionable. Instead of cost benefit analysis, evaluators often measure cost-effectiveness to determine the efficiency of a project or program. Cost effectiveness involves comparing costs of different approaches to achieving a given objective. Examples include the average cost of registering a voter or per unit cost of providing fertilizers to farmers.

### *3. Impact*

In evaluation parlance, impacts refer to results or effects that are caused by, or are attributable to, a project or program or policy. Impact evaluations usually focus on higher level effects of a project, program or policy that occur in the medium or long term. For example, have farmers' incomes increased as a result of an agricultural development project or fertility rates declined due to family planning? Such effects can be intended or unintended, positive or negative. Ideally, impact evaluations should be based on baseline data for both an intervention group and a comparison or control group, and include a second and third round of data after the intervention. Unfortunately, such data are not gathered in most foreign assistance interventions for various reasons, and evaluators have to use counterfactual reasoning (a hypothetical statement of what would have happened had the program not been implemented) to assess impacts.

### *4. Sustainability*

Sustainability refers to the continuation of an intervention's services and benefits after foreign assistance ends. Three dimensions of sustainability – financial, institutional and environmental – can be examined in evaluations. Financial sustainability indicates the capacity of an agency or organization assisted by a project or program to be financially self-sufficient, either through revenue-generating activities or through substitution of other public, private, or donor sources of funding. Institutional sustainability refers to the supported organization's capacity to manage its operations independently. Finally, environmental sustainability refers to the capacity of an intervention's services and benefits to survive in the changed or changing environment.

### *5. Relevance*

Evaluations also examine the continued relevance of the intervention's objectives and approach in light of changing development problems, policies, or priorities. The political, economic and institutional environments in which projects and programs are designed and implemented tend to change over time. Some of the changes can have major consequences for a project or program. For example, if a host government decides to withdraw food subsidies this would undoubtedly affect on-going agricultural development interventions, or a military coup will have implications for existing democracy programs. Under these circumstances, an evaluation can be undertaken to



find out if ongoing interventions remain relevant in the changed circumstances and if not, what changes can be made to make them relevant and useful.

## **1.2 OBJECTIVES OF FOREIGN ASSISTANCE EVALUATIONS**

Well designed and empirically grounded evaluations of foreign assistance should promote the following four distinct but inter-related objectives:

### ***1.2.1 Improved Performance***

Most evaluations are undertaken to improve the performance of a project or program by generating knowledge, information and recommendations which can be used by their managers. For example, a mid-term evaluation of a civil society program in an African country can tell program managers why it is not reaching the targeted civil society organizations, what concerns and reservations these organizations have about the program, and finally what can be done to solve the problem. Such information will help managers of projects or programs to make mid-course corrections to improve future performance.

### ***1.2.2 Accountability***

Well designed evaluations also provide accountability information for the USG resources spent on foreign assistance. Together with the data generated by FACTS and other information systems, they can find out if USG policies, strategies, programs or partners have achieved what they were supposed to achieve and if not, why not. They can also determine the cost effectiveness of assistance programs as well as the quality of their planning and implementation. Consequently, evaluation findings can further illuminate reports to different stakeholders – Office of Management and Budget, the White House, the Congress, the NGO community and, above all, the American taxpayers.

### ***1.2.3 Informed Decision-Making***

Evaluations also contribute to more informed decision-making. Officials administering foreign assistance programs have to make critical decisions affecting the present and future directions of foreign assistance. For example, a post/mission in a Rebuilding Country might have to decide whether it should continue to invest heavily in civil society; or the Office of the Director of U.S. Foreign Assistance might have to make a decision about the future allocation of resources to a developing country. In both cases, evaluations can aid informed decision-making. By commissioning an evaluation, concerned officials will be able to decide a future course of action on the basis of solid data and independent analysis rather than hunches or hearsay.

### *1.2.4 Lessons from Experience*

Finally, foreign assistance evaluations codify organizational and programming experiences thereby facilitating learning by experience. They generate knowledge which is fed into the development of new projects, programs, intervention strategies and policies. A simple example illustrates this. In the 1990s, the USG was heavily engaged in supporting post conflict elections in Rebuilding Countries. A set of seven evaluations on assistance for such elections, undertaken by USAID, provided several policy and operational lessons which profoundly affected subsequent USG assistance to post conflict elections. The reality is that a learning organization requires an effective evaluation system, which generates, stores and disseminates empirically grounded knowledge.

It should also be noted that there are legal and regulatory mandates for evaluation under the U.S. Code and Government Performance and Results Act (GPRA) of 1993. These requirements were further reinforced by OMB by introducing the Program Assessment and Rating Tool (PART) (See Annex 1)).

## **1.3 EVALUATION PRINCIPLES FOR FOREIGN ASSISTANCE**

There are certain principles which should inform all evaluations of foreign assistance. These principles are widely recognized by the international development community as well as by evaluation professionals.

### *1.3.1 Usefulness*

The first principle is usefulness. A critical distinction between academic research and an evaluation enterprise is the usability of the generated information. The purpose of academic research is to generate knowledge, establish cause and effect relationships, and contribute to theory building. Evaluation, on the other hand, seeks to gather information, ideas, lessons and recommendations that can be put to use by decision-makers and managers. If an evaluation fails to generate usable information, the time and resources spent on an evaluation are not justified. In practice, this principle implies that all foreign assistance evaluations should contribute to one or more of the following objectives (a) improving performance (b) promoting accountability, (c) fostering informed decision-making, and (d) learning by experience.

The usefulness principle imposes obligations both on those who commission an evaluation and those who conduct it. Evaluation managers should write comprehensive and clear scopes of work (SOWs) that precisely identify evaluation questions and allocate sufficient resources. They should constantly keep in touch with the evaluators to keep them on the right track. The evaluators' obligation is to answer questions on the

basis of empirical data and rigorous analysis and make recommendations that are actionable and relevant.

### *1.3.2 Credibility*

The second principle is credibility. The credibility of an evaluation largely depends on its methodological rigor and quality of data. It is incumbent on evaluators to tap all sources of information, use appropriate research methods and follow norms of scientific inquiry. Since reliable and comparable quantitative data are not always available, evaluators should use multiple methodological approaches to fill information gaps.

Methodological rigor should not be confused with quantitative methods. Qualitative methods are necessary to generate ideas, answer the questions of why and how, and derive practical recommendations. Both modes of data collection and analysis – qualitative and quantitative- can be rigorous or weak. What is generally needed in foreign assistance evaluations is to judiciously use multiple methods and to strictly adhere to scientific norms of inquiry. When possible, foreign assistance evaluation should, therefore, use both methods.

### *1.3.3 Independence*

The third principle is independence which means that evaluators should be free from any official interference. If evaluators are hired by an organizational unit which manages the assistance program being evaluated, it is possible that they will be under its influence therefore constrained in telling the truth as they see it. Lack of independence has sometimes been seen as problem with foreign assistance evaluations.

The independence of evaluations can be ensured in three ways. One approach is to ensure that evaluations are commissioned by a management unit which is separate from and superior to the one which manages the program. For example, a project and program evaluation should be managed by a post/mission office and not by the manager of the concerned project or program. Whereas, country evaluations should be conducted by a regional bureau, a central evaluation unit of the Agency, or the Office of the Director of U.S. Foreign Assistance, and not by the country mission itself. Another approach is to entrust the responsibility of conducting evaluations to an outside research or evaluation organization, which is not accountable to the managers of the program. Outside research and evaluation organizations pride themselves on their objectivity and independence and are in a position to resist the influence of a funding entity. The third and last option is to foster an organizational culture that emphasizes the need for rigorous and independent evaluations. Managers learn that it is in their own interest that the findings and recommendations reflect objective data and the considered judgment of the evaluators. They should not be afraid of criticisms and

negative findings; as such findings often produce the greatest value added in informing mid-course corrections and future program designs and thereby enhancing the chances for program success

#### ***1.3.4 Capacity Building***

The last, though not the least, governing principle is evaluation capacity building in partner countries. Capacity building is of prime importance in the context of foreign assistance, as the ultimate goal of all foreign assistance programs is to enable a country to solve its own problems. Capacity building requires that evaluations should contribute to the diffusion of evaluation skills and expertise and their institutionalization in host countries. This can be accomplished in many ways. One simple way is to include at least one local researcher in each evaluation conducted by posts/missions. Evaluation managers should encourage the evaluation teams to include host country nationals. Another approach is to award evaluation contracts to local firms and institutions. Even when evaluation cannot be locally contracted, local firms can often be sub-contracted to conduct surveys, focus groups and community interviews. Such approaches will not only build local institutional capacity but also provide an insider's perspective, which is often missing in foreign assistance evaluations. Still another way is to allocate funds to organizations to build local capacities in monitoring and evaluation. Many bilateral and multilateral agencies have been providing assistance to research institutions and universities to develop expertise in statistical data collection and analysis, monitoring, survey research and evaluation. However, much more needs to be done.

### **1.4 FOREIGN ASSISTANCE EVALUATION CATEGORIES**

Depending on their perceived needs and requirements, different agencies, bureaus and operating units engaged in administering foreign assistance - can undertake a variety of evaluations. A few generic categories for evaluations are identified here. The list is illustrative and not exhaustive.

#### ***1.4.1 Project Performance Evaluations***

The most basic evaluations are project performance evaluations which are also labeled as process, mid-term or simply project evaluations. These evaluations focus on the performance of a project and examine its implementation, inputs, outputs and outcomes/results. A project is defined as an individually planned undertaking designed to achieve specific objectives within a given budget and time frame. Performance evaluations tend to grapple with a range of simple but important questions: Did the project take off as planned? What problems and challenges, if any, did it face? Is it being effectively managed? Is it providing planned goods and services

in a timely fashion? If not, why not? Were the original cost estimates about the project realistic? Have the host country policies, economic environment or political circumstance changed? Will the project be able to meet its targets? What are its intermediary effects and impacts? What can be done to improve its performance and impacts? The overarching objective of these evaluations is to answer such questions to improve performance of an ongoing project.

Because they focus on a single intervention, project performance evaluations are relatively easy to undertake. Usually, a team of two or three evaluators can complete them within three to four weeks. In countries which have acquired evaluation expertise, project evaluations should be commissioned to local firms or research organizations, provided no sensitive information is involved. If the team is led by expatriate experts, a local evaluator/researcher should be included in the evaluation team whenever possible.

Most of the information for such evaluations can be gathered through reviews of project and program documents; interviews with project managers, host country officials and other stakeholders; mini-surveys, focus group discussions and meetings with beneficiaries. Such evaluations usually do not require sophisticated research designs; however, if quantitative data are to be collected from project beneficiaries, random sampling procedures may be needed.

### ***1.4.2 Project Impact Evaluations***

Project impact evaluations differ from performance evaluations in that the focus is on the effects and impacts of a project. Also known as ex-post evaluations, these are usually, though not always, conducted either when a project is likely to end or has ended. Such evaluations should cover both intended and unintended effects of a project--which can be positive and negative.

Although the situation differs from project to project depending upon its objective, nature, focus, duration and the overall institutional environment, impact evaluations should generally be designed to answer questions such as: Did the project achieve its stated objectives? What changes were observed in targeted populations, organizations or policies during the life span of project? To what extent can the observed changes be attributed to the project? How did the project affect the recipient country organizations which were involved in it? Were there unintended effects of the project? Were they positive or negative? What factors explain the intended and unintended impacts of the project?

It is difficult to measure precise impacts of complex foreign assistance projects, especially in the absence of baseline data. Even if baselines data are available, they are

often not sufficient to ensure that the observed changes can be attributed to the intervention if similar data are not available for a control group which did not receive an intervention's goods and services. However, a major problem for assessing impact is that in most cases, baseline data are not available for the project, much less for the control group. Only in a small number of cases, particularly in the health and education sectors, are baseline data gathered. Even then, their reliability is often questionable. Under these circumstances, evaluators are forced to build a counterfactual (what if?) case for assessing, not measuring, the changes that have occurred. Foreign assistance agencies are now increasing efforts to encourage operating units to collect baseline data so that more reliable and valid lessons can be drawn about project impacts. Such efforts should be strongly supported.

### **1.4.3 Program Evaluations**

These evaluations, which focus on a program area in a country, are complex research enterprises, as they cover multiple interventions, undertaken at different times, with different partners often with different sources of funding. However, they are necessary to give a total picture of USG programs – their achievements, impacts, failures and challenges – in a thematic area.

For example a post/mission may commission such an evaluation to examine the implementation of its various democracy promotion projects funded by the State Department and USAID and implemented by different NGOs and private sector firms.

Program evaluations require careful planning and preparation. Their objectives should be clearly defined and they should focus on a limited range of questions. The sampled projects and activities should be representative of all interventions in the sector or sub-sectors. Usually a team of 3 to 4 evaluators with strong research backgrounds and subject matter expertise are needed to conduct them. Depending upon the size of the programs, their history and geographical spread, four to six weeks should be budgeted for such evaluations.

As program evaluations require considerable investment of time and resources, posts/missions should use them selectively. They are particularly useful when a post/mission has to reorganize and reorient its existing programs in the face of new opportunities or challenges. For example, a country is reverting back to its old authoritarian ways, and the post/mission has to reorient its democracy assistance activities to make them more relevant to new political realities. An evaluation of its democracy promotion interventions can outline different programming options for the post/mission's consideration.

### ***1.4.4 Global/Regional Impact Evaluations***

Global/regional impact evaluations, which are designed to examine the impacts of USG assistance to a group of countries in a program area, are more complex than single country program evaluations. Evaluation teams have to construct a theory of development change and marshal all kinds of data and information-- often of varying quality--to assess the impacts. They have to visit a sample of countries and conduct on site studies. Their work is greatly complicated by an absence of baseline information.

Like project impact evaluations, global/regional impact evaluations focus not only on the intended results but also on impacts which were not anticipated by program planners. Often, unintended effects can be as significant as the planned effects. For example, many micro-finance programs which were designed primarily to promote economic entrepreneurship among poor, marginalized women, have also raised social and political awareness among beneficiaries leading to their increased participation in social and political affairs. Consequently, a comprehensive impact evaluation of a microfinance program should not only assess its effects on the income of women participants but also on their identities, status in the family and even social and political participation.

Global/regional impact evaluations are necessary to explore new directions for a program area. They are also useful to decide about the future of an ongoing assistance program. For example, if an agency/bureau has to decide if it should continue to allocate resources to improve the enabling environment for trade and investment, a global impact evaluation can undoubtedly help. By generating information about the overall results of trade and investment enabling projects in different countries and by identifying the factors that affected their performance and impacts, the evaluation team should be able to craft sensible recommendations for the concerned agency/bureau. Global/regional impact evaluations are particularly recommended when the theory behind the program is not empirically grounded, there is limited and questionable information about the effects of programs, and/or the USG has to make strategic choices about future investments.

### ***1.4.5. USG Country Assistance Evaluations***

Such evaluations focus on the entire USG foreign assistance funding to a country. Examples of the issues country evaluations can explore include (a) the country's progress or lack of progress towards achieving transformational diplomacy goals; (b) coordination among USG entities administering assistance programs; (c) the impact of USG assistance on selected sectors which received or continue to receive most assistance, and (d) the continued relevance of USG assistance in changed economic, political and strategic contexts.

As they cover the activities of the State Department, USAID and other USG entities, the Office of the Director of U.S. Foreign Assistance is perhaps the most suitable entity to initiate and manage such evaluations. The other alternative is to undertake them as collaborative endeavors involving the concerned bureaus/operating units of the USG. Country evaluations are most appropriate for the societies in which the USG has made or is making large investments of its foreign assistance resources.

Such evaluations require strategic thinking and methodological triangulation. In addition to an intensive analysis of the available data, documents and records, evaluators should conduct in-depth interviews with experts and policymakers both in Washington and in the host country. When the purpose is to assess the impact of USG assistance, evaluations would also require analysis of available macro-level data from different sectors. USAID has conducted such country evaluations, but their number has been very small. Moreover, they focused on USAID funded activities alone and not on all USG assistance to a country. Typically, the best approach is to commission such evaluations through outside research organizations, which have substantive and technical expertise in the area.

### ***1.4.6 Organizational Evaluations***

Organizational evaluations focus on management processes of an organization or of its component units. Their purpose is to assess the way an organizational entity manages foreign assistance programs. They also examine if they conform to foreign policy objectives, laws, regulations and policies. Such evaluations are desirable to improve the management and delivery of foreign assistance.

Organizational performance evaluations should be carried out when there is a need to examine topics such as (a) the performance of an organizational unit; (b) procedures for planning and implementing foreign assistance programs; (c) duplication of functions among different organizational entities, and (d) cost-sharing. Posts/missions, bureaus, or agencies can commission such evaluations.

### ***1.4.7 Special Evaluation Studies***

Such evaluations are driven by the specific information needs of the State Department, USAID, bureaus, post/missions, and partner organizations. They may be undertaken when (a) a key decision has to be made and available information is inadequate; (b) there are major implementation problems that should be addressed, and (c) major stakeholders – the Secretary of the State, OMB, White House and Congress – need empirically grounded information that is not available from routine sources.



## 1.5 EVALUATION MODALITIES

The following four modalities for evaluations of foreign assistance programs are defined on the basis of who conducts them.

### *1.5.1 Internal Evaluations*

Internal evaluations are conducted by an operating unit to assess the progress of its activities, identify problems and bottlenecks, and find ways to improve its performance and impacts. These are management exercises which are usually done by outside experts. If conducted freely and professionally, internal evaluations are extremely useful. For example, an internal evaluation of a civil society project can inform its manager about its achievements and failures and generate ideas and recommendations to improve its performance. The main limitations of internal evaluations are that they can be self-serving and lack credibility. Since they are commissioned by the management, evaluators are often reluctant to criticize management's performance. As a result, outsiders cannot be sure about their objectivity. On the other hand, since managers commission them, they are more likely to accept the findings and recommendations. Most of the evaluations undertaken by foreign assistance agencies follow the modality of internal evaluations.

### *1.5.2 External (Independent) Evaluations*

These evaluations are administered by an operational unit which is separate from and/or superior to the unit managing the project or program and are carried out by an independent outside group. For example, the evaluation of a project to provide technical assistance and support for the election commission in a country would be managed by the office of the mission director, or by a Washington Bureau, and not by the democracy officer who is in-charge of such assistance. External evaluations tend to be more objective and credible, as evaluators are not responsible to managers of the project or program being evaluated. Evaluators enjoy greater freedom than they do in internal evaluations, and are able to express their criticisms more freely and candidly. Often, such evaluations are more formal than internal evaluations, as evaluators have to defend their findings before the managers and staff of the concerned project or program.

There is a widespread perception among foreign assistance stakeholders – particularly the Congress, OMB and the academic community – that most evaluations conducted by foreign assistance agencies are neither rigorous nor independent. Such perceptions can be largely dispelled by promoting external evaluation.

### ***1.5.3 Participatory Evaluations***

Such evaluations involve active participation of multiple stakeholders, which include beneficiaries, the implementing organization and the operating unit, in planning, data collection, analysis, reporting, dissemination and follow-up actions. A common modality involves the implementing organization collecting background material and circulating it among the stakeholders. These stakeholders analyze the material and explore its implications in a workshop or a series of workshops. Findings and recommendations are formulated by a panel. USAID missions have conducted such evaluations mostly on an informal basis.

Participatory evaluations have obvious advantages. They enable managers of operating units to listen to, and respond to, stakeholders. Face to face interactions facilitate better understanding of the workings of a project or program and its achievements and problems. Participants often come up with new ideas for solving problems or improving performance. As managers themselves participate in the evaluation process, they are inclined to use resulting information and recommendations. However, participatory evaluations have many limitations. Such evaluations tend to be less objective because participants have vested interests which they articulate and defend in such workshops. Moreover, they are less useful in addressing complex technical issues, which may require specialized technical expertise. Still another limitation is that although they may generate useful information, their credibility is limited because of their less formal nature.

### ***1.5.4 Collaborative (Joint) Evaluations***

Collaborative evaluations are conducted jointly by more than one office, agency, or partner. For example, a collaborative or joint evaluation might be conducted by a team comprising staff from a USAID mission, the World Bank, the host country, and an NGO.

There are several advantages to collaborative evaluations of USG foreign assistance programs. They facilitate mutual learning among partnering organizations. The burden for conducting evaluations is shared among collaborating entities as a single evaluation is undertaken in place of multiple evaluations by collaborating organizations. As more than one organization is involved, evaluators usually enjoy greater flexibility and independence than in evaluations commissioned by a single organization. Finally, as a result of the pooling of resources, more intensive and systematic data collection and analysis can be undertaken which can enhance both the legitimacy and credibility of evaluation findings and conclusions.

On the negative side, collaborative evaluations can be time consuming. For example, collaborating organizations may find it difficult to agree on a common terms of reference (TOR) for an evaluation. They tend to differ about evaluation designs and data collection methodologies, and often require time to evolve a consensus around these issues. Collaborative evaluations also require a strong management team which can keep in touch with different partners and reconcile their concerns and expectations. Unless carefully planned, the findings of collaborative evaluations may not always meet the needs and requirements of all partners. Concerned bureaus and operating units should weigh these factors before launching a collaborative evaluation.

The choice of modality for an evaluation depends on perceived needs, resources and circumstances. For example, if objectivity and credibility are key requirements, an external evaluation may be the appropriate choice, whereas if stakeholder ownership and acting on findings are priorities, more collaborative or participatory approaches are usually better.

### **1.6 INSTITUTIONAL RESPONSIBILITIES FOR EVALUATION**

#### ***1.6.1 Post/Mission***

The responsibility for conducting project/program evaluations rests primarily with the post/mission and its operating units and partners. Each post/mission may designate a monitoring and evaluation officer to supervise the planning and conduct of its evaluations and the utilization of their findings.

Each post/mission may perform the following functions:

1. Ensure that all its major interventions are evaluated; their problems, achievements and impacts analyzed; and reliable and relevant lessons learned are developed to guide future interventions.
2. Allocate sufficient funds to monitor and evaluate each project/program. As a general rule, 5-10% of the funds for a program/project may be allocated for monitoring and evaluation purposes, depending on the size of the program
3. Ensure the integrity and independence of evaluations. All evaluations should follow the principles identified earlier in these guidelines.
4. Work with host governmental and non-governmental organizations to improve their expertise and capacities to monitor and evaluate development interventions. This is necessary to reduce the cost of evaluations and to build institutional capacity in a country.

5. Facilitate the dissemination and utilization of the findings and recommendations of evaluations within and outside posts/missions, including to Washington-based repositories. Information on such evaluations should be included in the Operating Units' (OU) annual Performance Plan and Reports (PPR).

### **1.6.2 Bureaus**

Many State and USAID bureaus have evaluation offices which are generally independent of central evaluation units. Depending upon their size, resources and mandate, the bureau-specific evaluation office may perform the following functions:

1. Formulate specific procedures and modalities for monitoring and evaluation of programs relevant to the bureau's objectives and operations.
2. Conduct evaluations of the bureaus' projects and programs, using rigorous data collection and analysis methods.
3. If and when necessary, arrange for technical assistance to operating units in the field for undertaking empirically grounded evaluations in a timely fashion.
4. Facilitate the dissemination and utilization of evaluation findings so that the bureaus' other projects and programs can also profit from past experiences.
5. Document the results of evaluations in the Bureaus' annual PPRs

### **1.6.3 Central Evaluation Units**

It is expected that both the State and USAID will have central evaluation units over time. When operational, the central evaluation units in State and USAID may perform the following functions:

1. Provide overall guidelines, standards and procedures for evaluations of foreign assistance programs.
2. Work with bureaus, posts/missions and other organizations to ensure compliance with evaluation procedures and guidelines in the field.
3. Arrange for technical assistance to facilitate empirically grounded evaluations in a timely fashion.
4. Arrange for training of foreign assistance staff in monitoring and evaluation.
5. Conduct country, regional and global evaluations of agency/bureaus' projects and programs. Such evaluations may involve sending teams to the field to (a) identify any bottlenecks in specific sets of projects and programs; (b)

- document the outcomes and impacts of the agency/bureaus' interventions, and (c) draw lessons from experience to develop future innovative projects/programs.
6. Maintain evaluation documents, records and other relevant data and information in an easily accessible web-site format for the use of evaluation staff and to inform decisions by managers.
  7. Facilitate the utilization of evaluation findings.
  8. Represent the State Department and USAID in national and international forums and networks on evaluations.

### ***1.6.4 Office of the Director of U.S. Foreign Assistance***

The Office of the Director of U.S. Foreign Assistance plays a coordinating and supporting role in the evaluation of foreign assistance projects and programs. It may undertake the following functions:

1. Coordinate with other agencies to establish sound policies, procedures, and standards for monitoring and evaluating foreign assistance programs; to promote the conduct of appropriate evaluations to assess and improve program performance; to strengthen staff skills in monitoring and evaluation; and to document and disseminate lessons learned.
2. Support USAID and State efforts to strengthen central monitoring and evaluation functions. As appropriate, initiates evaluations of cross-cutting programs to assist with policy and program decisions, improve performance and document lessons learned.
3. If and when necessary, initiates evaluations of cross-cutting programs independently or in cooperation with USAID, State Bureaus, and other units of the U.S. Government. These evaluations will differ from the evaluations conducted by USAID, State bureaus and other offices in two respects. First, they will focus on programs involving multiple foreign assistance operating units. For example, in a country or region, where counter-narcotics programs are administered both by USAID and State Department, each might properly evaluate its own interventions independently. However, the DFA scope would be to evaluate the entire USG counter-narcotics assistance to a country or region or world wide. Second, as compared to the evaluations done by other agencies and bureaus, its evaluations will be more strategic in nature and scope.

A few examples of evaluation topics which the Office of the Director of U.S. Foreign Assistance may undertake include: (a) strengths and weaknesses of USG assistance programs in a major recipient country or group of countries; (b) impact evaluations of USG foreign assistance in a program area in a country or region; (c) regional and global evaluations of specific categories of projects and programs; (d) evaluations or special studies of existing policies and procedures for delivering foreign assistance, and (e) evaluations or special studies of the topics, problems or issues which the Director of U.S. Foreign Assistance wants to examine in order to pursue transformational diplomacy goals.

4. Analyze Performance Plans and Reports to assess the sufficiency of level of effort in evaluation and to inform the development of future evaluation plans.
5. Maintain contacts with the evaluation offices of bilateral and multilateral donor agencies and professional organizations. Represent U.S. Foreign Assistance in international meetings and seminars on monitoring and evaluation and related topics.

## **PART 2: MANAGING THE EVALUATION PROCESS**

### **2.1 PLANNING EVALUATIONS**

#### ***2.1.1 Specifying Objectives***

The first step in planning an evaluation is to decide if an evaluation will be useful. Once decided the manager should then proceed to define the objectives and audience for the evaluation. Specification of objectives helps in sharpening the focus of an evaluation and formulating evaluation questions. It also makes it easier for evaluators to come up with relevant findings and recommendations.

An evaluation's objectives should be stated in such a way that they also specify how the information generated by the evaluation will be utilized. It is not enough to state that the objective of evaluation is to examine the performance of a family planning project or assess its impacts. It is also necessary to specify, whenever possible, how the findings will be used to improve the performance of the project or how the lessons of an ex-post evaluation will feed into planning new projects on family planning in Africa.

To determine the objectives of evaluations, managers should deliberate over questions such as: who needs evaluation findings and recommendations, why do they need them, and how will they use them? Evaluation managers should also determine when evaluation findings and recommendations will be used. For example, before deciding to commission an evaluation of its civil society assistance program, the evaluation manager (who is most likely the supervisor of democracy programs) in a post/mission should seriously examine his/her own information needs and expectations from the evaluation. Does he/she want to know about the implementation problems faced by the program? Does he/she want to find out if the program is likely to meet its targets over time? Does he/she require information about the sustainability of the program when funding stops? Does he/she need to find out about the overall impacts of the program? By focusing on such questions and discussing them with concerned staff, the evaluation manager can specify the objectives of the evaluation.

Evaluation managers should also consider the needs of other stakeholders – implementing organizations, the host country or its own bureau in Washington -who might be interested in the evaluation and could benefit from it. They should therefore consult them while defining the evaluation's objectives. For example, if the democracy officer in a post/mission plans an evaluation of its media assistance project, which provides short and medium term training to local journalists, he/she should consult with the implementing partners, staff of the local training institutions, media outlets at which trainees are employed and other concerned organizations to solicit their views

and suggestions.

### ***2.1.2 Formulating Study Questions***

Evaluation questions logically follow from an evaluation's objectives. For example, if the purpose is to find out why farmers are not adopting technical packages in an agribusiness project, questions will be different from when the objective is to assess the project's impacts on farmers' income or the sustainability of the marketing organization established by the project.

A single evaluation cannot answer many questions unless the time and resources are abundant, which is rarely the case. Therefore the number of questions should be limited. Three criteria should be used to select evaluation questions. First, questions should reflect management priorities. For example, if the post/mission wants to know about the sustainability of a project after USG funding stops, questions should primarily pertain to sustainability. Second, only those questions which require the use of empirical data to answer them should be included. Evaluators are supposed to answer questions on the basis of hard evidence (both quantitative and qualitative) and not subjective opinions. Third, the data and information to answer questions must be able to be gathered within the parameters of the given time and resources. Thus, questions which may require a long time or very large financial resources should be omitted if the concerned operating unit lacks such time and resources.

Those stakeholders who are likely to utilize the evaluation's findings should also be consulted in framing questions.

### ***2.1.3 Evaluation Design and Methodology***

Evaluation managers are not expected to develop detailed evaluation designs or methodologies. This is a task which should be left to evaluation teams.

It should be noted here that most evaluations of foreign assistance programs do not require complex "experimental" and "quasi-experimental" designs including randomized control trials. Such designs are usually not appropriate for examining management and implementation issues. They are also not relevant to institution building projects and programs. However, "experimental" and "quasi-experimental" designs are desirable for measuring the effects or impacts of a project or program.

Evaluation managers should pay particular attention to two issues (a) the quality and credibility of the data and information they need, and (b) the type of data collection methods (case studies, sample survey, comparative evaluation design, rapid appraisal methods, analysis of existing data, participatory workshop, and the like) that can generate the needed information. Different data collection methods have distinct features that make them either more or less appropriate for answering particular types



of questions. For example, if the question is what percentage of the population has become aware of the menace of HIV/AIDS in a country, then a sample survey would be most appropriate. If, by contrast, the issue is why public education programs are not succeeding in disseminating information about HIV/AIDS, qualitative methods such as focus group discussions, key informant interviews or group meetings would be a better choice.

One simple way to identify appropriate data collection methods is to examine past evaluations of similar projects and programs. Evaluation managers now have access to thousands of evaluations as well as other resources on the USAID/ DEC or State Diplopoedia web site. They can also review thousands of evaluation reports which are posted on Web-sites by various international agencies (See Annex 2 for more information on such resources). In addition, they can also consult with experts to gain additional information about the suitability of various data collection methods.

### ***2.1.4 Assessing Evaluability***

In large and complex evaluations, particularly country, multi-sector program and regional and global evaluations, it is advisable to undertake an evaluability assessment. Such an assessment examines the feasibility of the proposed evaluation with reference to three considerations. First, whether the evaluation objectives and questions are plausible and relevant data and information can be gathered within the given time and resources. The feasibility of answering questions of major interest to the commissioning unit or stakeholders should be reviewed in relation to the investment of available time and resources, and a recommendation on feasibility made. In other cases, the assessment might find that the evaluation team simply would not be able to gain access to the required data. For example, if no baseline data have been gathered, the evaluator cannot precisely measure the changes which can be attributed to an intervention.

The second consideration concerns the consent and cooperation of the stakeholders, particularly the host country and partnering institutions. Evaluation managers should consult with major stakeholders, seek their cooperation and address their doubts and concerns when the evaluation is being designed.

Finally, the overall cost of the evaluation should be considered. Ideally, the expected benefits from an evaluation should be at least of equal value to the resources expended on it. While it is not easy to quantify the potential benefits of an evaluation, evaluation managers should examine the cost effectiveness of the proposed evaluation and the potential use of its findings and recommendations.

## 2.2 CONTRACTING AND FINALIZING EVALUATION DESIGN

### 2.2.1 *Preparing the Scope of Work*

The next logical step is to prepare a scope of work (SOW) for the evaluation. A SOW is a blue print, a coherent plan of action. Past experience indicates that a well-written and well-thought-out SOW ensures that the evaluation will meet the information needs of the commissioning unit. Moreover, it reduces possible misunderstandings which might arise between evaluation managers and evaluation teams.

The time and resources needed to prepare a SOW vary from evaluation to evaluation. If the evaluation manager is familiar with the concerned project and the scope and objectives of the evaluation are limited, preparing a scope of work is not difficult and time consuming. For example, a USAID Mission economist who supervises an agricultural marketing project can easily write a SOW for a mid-term evaluation of the project, as he/she is intimately familiar with it. On the other hand, if the evaluation manager is based at headquarters, and has to write a SOW for a multi-country evaluation of agricultural development programs, he/she will have to spend considerable time in reviewing relevant documents, interviewing staff and contractors working on the programs and contacting host governments and overseas missions to define objectives and write the evaluation questions. He/she may even need technical support to prepare the SOW.

An evaluation SOW should cover the following:

- Purpose of the evaluation – scope, audience and intended use of findings
- A brief history and current status of the project/program/ strategy to be evaluated
- Evaluation questions
- Available information and sources – performance data, FACTS data, previous evaluations
- Preliminary evaluation design and data collection methods
- Qualifications of evaluators and composition of evaluation team
- Desired evaluation reports
- Time schedule
- Budget

### 2.2.2 *Criteria for Selecting Evaluators*

USAID and State tend to contract out most of their evaluations. They select evaluation firms or individual evaluators through transparent procurement procedures, e.g. an evaluation Indefinite Quantity Contract (IQC) or General Services Administration Schedule, that are well-documented and need not be mentioned here. However, while reviewing evaluation proposals, technical review teams should pay particular attention

to the following five criteria:

First, Evaluation Skills: Reviewers should carefully examine the technical skills and experience of evaluators. Such skills include their training and experience in evaluation design and methodology. Ideally, evaluators should be able to develop evaluation designs; collect data and information using various data gathering methods; establish rapport with stakeholders; and write empirically grounded reports, with actionable recommendations. In addition, it is important that evaluators are able to work as a team. It is also strongly recommended that the Team Leader have a depth of experience in both evaluation and team management.

Second, Subject Matter Expertise: Knowledge of evaluation methodology is a necessary but not a sufficient condition for evaluating foreign assistance programs. Therefore, if the proposed evaluation focuses on HIV/AIDS, some evaluators must have a strong background in and understanding of the HIV/AIDS pandemic and of assistance projects and programs to deal with it. Similarly, if an evaluation is designed to examine the performance of an electoral assistance project, some evaluators must be familiar with the planning, holding and monitoring of elections, as well as the nature of international electoral assistance programs.

Third, Country Knowledge: Since foreign assistance evaluations focus on projects and programs in foreign countries, the knowledge and understanding of the host countries is generally necessary. Familiarity with the local language is especially necessary when the evaluation requires interviewing project or program beneficiaries. Other things being equal, preference should be given to evaluators who possess language proficiency, in-country work experience and a broad understanding of the host country context in which a project or program operates.

Fourth, Gender Mix and Gender Analysis Skills: Representativity on evaluation teams usually aids in data collection. For example, women evaluators are often more effective in soliciting information and ideas from women beneficiaries, who tend to be more comfortable with them rather than with male interviewers. The reverse is also true. When the target population is mixed, reviewing teams should pay particular attention to the gender mix of evaluation teams as well as the team's capacity to undertake gender analysis, and data disaggregation by sex if and when necessary.

Fifth, Conflict of Interest: Care should be taken that evaluators have no potential biases or vested interest in the evaluation outcomes. This requirement may limit the participation of those firms and evaluators who have a direct stake in the concerned project or program. This is necessary to ensure the independence and integrity of the evaluation.

### ***2.2.3 Revisiting the Scope of Work and Evaluation Proposal***

After an evaluation team is selected, managers should carefully review with the team the original scope of work as well as the proposal submitted by the selected firm or evaluator. Such reviews facilitate better communication between managers and evaluators and help improve the quality of evaluations. They enable managers to clarify evaluation objectives, questions and their expectations from evaluations. At the same time, they give evaluators an opportunity to explain their proposals. Such discussions should particularly focus on the following:

- Evaluation questions: Do questions need revisions? Should additional questions be added and/or existing ones deleted?
- Evaluation design: Managers and evaluators should explore options other than those proposed in the SOW, if necessary.
- Time and resources: Are the amount of time and level of effort provided for the evaluation or assessment sufficient?
- Nature and content of evaluation report(s), briefs and presentations

In the light of these discussions, the evaluation team should develop a comprehensive evaluation plan possibly through a team planning meeting.

### ***2.2.4 Finalizing Evaluation Design***

In many cases, it may be necessary for evaluation managers to separately review the proposed evaluation design with the assistance of technical experts. Such reviews should focus on the following:

- Variables and indicators: These should relate to the objectives of the evaluation, the nature of data required and the feasibility of data collection. The FACTS data base can be very helpful in selecting output indicators.
- Proposed intervention cases should be examined, particularly in evaluations which cover multi-projects and or multi-sites. For example, in a case of a multi-country impact evaluation, the design should identify the countries which will be included in the study. In the case of the evaluation of a program in a country, the task involves selection of sites which will be studied. Such selection should be based on well articulated, objective criteria.
- Sampling may be necessary to select examples or cases from the population units. Evaluation managers should examine the proposed sampling technique(s). The choice of sampling depends upon the evaluation questions, the nature of investigation techniques (quantitative versus qualitative) and how precise and representative the results need to be.

- Units of analysis and sources of information: Units might be individuals, families, farms, communities, clinics, water wells, immunization campaigns etc. The sources of information can be individuals, groups, organizations, records and documents and published or unpublished data.
- There are a broad range of structured approaches to collecting quantitative and qualitative information. These include sample surveys, mini-surveys, case studies, key informant interviews, focus groups, community interviews, site observation, participatory workshops, and syntheses of existing documents. Managers should review the proposed data collection methods and ensure that they are appropriate to answer the evaluation questions.
- Different data collection methods use different types of research instruments. Surveys employ structured questionnaires, site observation techniques use observation forms, focus groups use loosely structured guides to facilitate and record discussions. Constructing reliable and valid instruments - questionnaires, interview protocols and guidelines for site visits – is critical to gathering reliable and valid data. When possible, evaluation managers should ask for research instruments and discuss them with the evaluation team.

If managers have questions or doubts on any topic mentioned above, they should seek clarification, as once the evaluation design is finalized it is too late to make any major changes. Although they should seek clarification, managers should respect evaluators' technical expertise and resist the temptation to impose their own views and preferences. For example, the manager of a large micro-credit program should ask questions about the criteria to be used for selecting sites for data collection, but he/she should not suggest which sites should be included or excluded. Such insistence will bias the sample and may distort the evaluation findings.

### **2.3 MANAGING DATA COLLECTION AND ANALYSIS**

#### ***2.3.1 Data Collection***

The primary role of managers at this stage is to keep in touch with evaluators, provide help and assistance when needed, and ensure that the evaluators are following agreed upon work plans.

Evaluation managers should also familiarize themselves with the data collection methods which evaluators are using. These are as follows:

- Document reviews, using established synthesis techniques to extract and analyze findings and lessons from existing progress reports and other documents.

- Secondary analysis of existing statistical databases, project records, files, and surveys, using a Results/Logical Framework approach. This approach begins by conceptualizing the hypothesized linkages of how an intervention or series of interventions are supposed to result in progressively higher levels of results, and then searches for existing evidence-- data--to prove or disprove these hypotheses (e.g., links data on project inputs and outputs to data on effective use by the target population to changes in behavior, quality of life, economic trends, or other impacts).
- Primary data collection using the five “rapid appraisal” techniques. These are key informant interviews, focus groups, community interviews, direct observation and mini-surveys. Key informant interviews involve interviews with 15 to 35 individuals selected for their knowledge and understanding of the topic. They usually reflect diverse views. Interviews are qualitative, in-depth and semi-structured. Interview guides listing topics are used, but questions are framed during the interviews, using subtle probing techniques. Focus groups are designed to generate ideas, issues, interpretations and recommendations. In such groups, 8 to 12 participants discuss issues and experiences among themselves. A moderator introduces the topic, stimulates and focuses the discussion, and prevents domination of discussion by a few. Community interviews take place at public meetings open to all community members. Interaction is between the participants and the interviewer, who presides over the meeting and asks questions following a carefully prepared interview guide. In direct observation, observers record what they see and hear at a program site, using a detailed observation form. Observation may be of physical surroundings or of ongoing activities, processes or discussions. Mini-surveys involve interviews with 25 to 50 individuals, selected using probability or non-probability sampling techniques. Structured questionnaires are used that focus on a limited number of closed-ended questions.
- Time series data gathered from rigorous sample surveys or censuses. Such surveys are particularly useful when baseline data exist so that comparisons can be made.

An evaluation study might employ several of these data collection and analysis approaches. In fact, in most cases, evaluation teams should rely on a mix of data sources, using both secondary data and primary data collection techniques.

Often, unexpected problems arise during data collection which may require changes in data collection plans. For example, an evaluation team member becomes sick and has to leave the team; evaluators suddenly find that the planned sample surveys cannot be

undertaken within the stipulated time, or concerned government officials, who were to be interviewed, have been transferred and the new officials are not familiar with the project or program. Under these conditions, evaluation managers should provide necessary help and assistance to evaluators. They should deal with the problem with an open and flexible attitude.

### ***2.3.2 Data Analysis***

Evaluators analyze data to discern patterns, trends, or comparisons. Whether in quantitative or qualitative data analysis, evaluators should use well-established methods. Quantitative methods include use of descriptive statistics including measures of central tendency (such as mean, median, and mode), regression analysis and analysis of variance to test the existence of potential relationships. The analysis of qualitative data is a more difficult, iterative process. It involves selecting, focusing and transforming raw data into meaningful categories, conclusions and recommendations.

Evaluation managers generally do not have time to engage in data analysis. This is perhaps not desirable, either. Their participation may create the impression that they are not unbiased observers. Evaluators also will feel uncomfortable when evaluation managers want to be involved in analyzing the gathered data and information.

### ***2.3.3 Reporting Evaluation Results***

Often the details about the nature and format of evaluation report(s) are mentioned in the SOW. Although there are no hard and fast rules, most evaluation reports should cover the following:

- Executive Summary
- Purpose, scope, and audience of the evaluation:
- Evaluation questions;
- The context (relevant history, demography, socio-economic status, and basic political arrangements of the community, country, or region) in which the project, program or policy intervention was designed and implemented;
- Description of the project or program, its objectives, scope and activities;
- The conceptual framework underlying the project or program;
- Data collection: the unit of analysis, sample, type of data collected, quality of data, and analytic techniques used;

- Evaluation findings: Findings are the empirical facts gathered by an evaluation team about the performance, cost effectiveness, relevance, sustainability or impacts of an intervention, strategy, policy or organization;
- Conclusions: These represent the evaluators' interpretations and judgments based on findings and the empirical data gathered and analyzed;
- Recommendations: These refer to the proposed actions which the evaluation team recommends to management. These should be logically derived from findings and conclusions;
- Lessons learned: These are broader implications for similar programs in different settings or for future activities.

Evaluation reports should not be written as academic documents. Instead they should be written in a way that can be easily understood by intended audiences. Technical terms and jargon should be avoided or at least explained. It often makes sense to present the main findings and recommendations up-front in the evaluation summary, as many managers and stakeholders are not interested in details about data collection and analysis. While evaluations should report negative findings, they should be presented in a constructive manner as the main objective of most evaluations is not to apportion blame but to find solutions to the problems and challenges facing projects and programs. Findings, conclusions, recommendations and lessons should be succinct, distinguished from one another, and clearly identified in the report.

There is often a tension between the needs of management and technical staff. While managers prefer concise reports focused on main findings and recommendations, technical staffs usually want much more detail about the process of arriving at the findings. Thus they want to know more about methodology, sources of data and literature review preceding the evaluation. One solution is to have a short report of 10 to 15 pages followed by detailed annexes. The other alternative is to write a summary of the report that is oriented to the needs of management. Still another alternative is to prepare different reports for different audiences as discussed later.

## **2.4 REVIEW AND FOLLOW-UP**

### ***2.4.1 Review of the Draft Report***

Evaluation managers should carefully review draft evaluation reports against the Scope of Work. They should also share them with their colleagues, outside experts and concerned stakeholders to seek comments and suggestions. Particular attention should be given to the following topics and questions:



- **Data Collection:** Are the data and sources of data clearly presented in the main report or an annex? Were the data collection methods consistent with the approved research design?
- **Quality of Data:** What is the quality of data and information gathered by the team? Are there serious questions about their reliability and validity? Does the report mention relative strengths and weaknesses of the data obtained in a transparent manner?
- **Coverage of Evaluation Questions:** Are all evaluation questions answered in the report? Are the data and evidence presented clearly? Are alternative explanations of findings explicitly considered and explored?
- **Recommendations:** Do the recommendations directly follow from the evaluation's findings and the conclusions? Are they supported by sound analysis and reasoning? Are they "actionable," in the sense that they can be implemented by the USG and its partners in existing circumstances?
- **Lessons:** Does the report mention lessons which may be used in designing new projects and programs? Are they adequately explained?

Evaluation managers should consolidate their own comments and those received from others in a single document. This is necessary, as they may receive comments and suggestions which are not compatible. The comments can be communicated both orally and in writing.

However, it is important that evaluation managers not try to influence evaluation findings and recommendations, which will compromise the integrity of the evaluation.

### ***2.4.2 Forms of Dissemination***

Evaluation findings and recommendations can be disseminated in many ways. The most common practice is to distribute the entire report. In the past, the practice was to send a printed version, but now it can be expeditiously done electronically, which saves time and resources.

Another option is to distribute a summary of the report to interested parties and stakeholders. A six to eight page summary tends to evoke a more positive response than the entire evaluation report, as managers and technical staff can quickly read them. The experience of USAID's former Center for Development Information and Evaluation (CDIE) confirms this. The Center used to publish a Highlight for each of its major program evaluations. These six to eight page Highlights were very popular with USAID staff and other development practitioners.

The third option is to prepare and disseminate evaluation briefs. Such briefs, not more than of one or two pages, succinctly summarize findings and recommendations. They are usually most effective in reaching out to senior officials. If potential readers are interested, they can ask for the entire report. Probably, the best course is to send evaluation briefs in the form of an email letter with the entire report as an annex.

In addition to distributing reports in written form, evaluation managers should encourage oral presentations by evaluation teams. Because they provide an opportunity for personal interactions, such presentations are more effective in communicating evaluation findings and recommendations than are written reports. Stakeholders and technical staff have an opportunity to ask questions about data, findings and recommendations. Evaluators, on the other hand, can offer necessary clarifications and further expound on their findings and recommendations. By creating a forum for discussion among managers and decision-makers, oral briefings often generate momentum for action.

Finally, USAID and State staff should include information on their planned completed evaluations in the annual Performance Plan and Reports in FACTS. This information needs to be accessible to agency managers and decision-makers, and is useful in preparing reports for key stakeholders, such as Congress.

### ***2.4.3 Follow-up***

As a follow-up to evaluations, the following steps should be taken:

First, senior managers responsible for the evaluated project/ program/policy should examine the evaluation findings and recommendations in-depth and explore their implications. For example, if an USAID mission commissioned an evaluation to examine the performance of its ongoing program on legislative strengthening, the mission director, program officer and democracy officer and the staff of implementing partners should discuss the findings and recommendations. They should determine which recommendations can be implemented. They should also identify management and/or program actions to be taken to implement those recommendations.

Second, the senior officials should assign clear responsibility for the completion of each set of actions. For example, in the case of a legislative strengthening program, if the evaluation has recommended revamping its training course for the newly elected members of the national parliament, the mission director/DG officer may ask the program director/chief of party to revise the training curriculum, incorporating the suggestions of the evaluation team by a specified date. He/she should also monitor progress in implementing recommendations.

Third, evaluation managers should share evaluation findings and recommendations with concerned organizational units, implementing partners and other stakeholders to promote learning. The extent of sharing, however, largely depends upon the scope and focus of the evaluation. For example, the dissemination of the findings of a mid-term evaluation of an agricultural project is likely to be limited to the staff of the operating unit, implementation partners and host government, if necessary. On the other hand, a program evaluation of public-private partnership may be distributed more widely. In all cases however the report should be submitted to the USAID/DEC or the State Department Diplopedia site. On the other hand, evaluation findings and recommendations of a global program on human trafficking should be distributed widely. A host of regional bureaus, implementing partners, consulting firms, NGOs, international donors and even Congressional committees might be interested in the findings.

Evaluation managers should exercise good judgment about with whom to share evaluation reports. However, the operating principle is that unless there are compelling reasons not to do so--such as that the document contains political or procurement sensitive materials--wider dissemination is desirable. If possible, politically and procurement sensitive sections can be deleted from the document so that a wider audience within and outside the USG can learn from the evaluation.

USAID requires copies of all final evaluation reports should be submitted in electronic form to USAID's Development Experience Clearinghouse (DEC). State Department is also implementing such a requirement, where reports should be submitted to the Diplopedia website (<http://diplopedia.state.gov>).

## **ANNEX 1: LEGAL AND REGULATORY MANDATE FOR EVALUATION**

U.S. law and legislation passed by the U.S. Congress has given USAID and the State Department a strong mandate to undertake evaluations.

### **Foreign Service Act of 1961**

Ever since its inception in 1961, Foreign Service Act has required USAID to place high importance on evaluations. The Foreign Service Act of 1961 required USAID “. . . to carry out a program of research into, and evaluation of, the process of economic development . . .”

### **U.S. Code**

The United States Code is the codification by subject matter of the general and permanent laws of the United States (<http://www.gpoaccess.gov/uscode/index.html>). The U.S. Code requires U.S. Agencies to undertake a systematic review of their operations, including the use of evaluations, to determine “the degree of efficiency and economy in the operation of the agency's activities, functions, or organization units.”

For instance, the U.S. Code (Title 5, Part I, Chapter 3, USC Sec. 305) requires that:

- (a) “Under regulations prescribed and administered by the President, each agency shall review systematically the operations of each of its activities, functions, or organization units, on a continuing basis.
- (b) The purpose of the reviews includes--
  - (1) determining the degree of efficiency and economy in the operation of the agency's activities, functions, or organization units;
  - (2) identifying the units that are outstanding in those respects”

### **Government Performance and Results Act of 1993**

In 1993, the U.S. Congress passed the Government Performance and Results Act (GPR) in order “to provide for the establishment of strategic planning and performance measurement in the Federal Government, and for other purposes.” Among other measures, the GPR passed into law three new requirements that relate directly to performance management and define and strengthen mandate to evaluate.

#### **Strategic Plan**

Like other agencies, State Department and USAID are required to produce a strategic plan containing “a description of the **program evaluations** used in establishing or revising general goals and objectives, with a schedule for future program **evaluations.**” (Emphasis added)

The GPR first required all Agencies undertake every three years, a five-year strategic plan for program activities. “Such plan shall contain:

1. A comprehensive mission statement covering the major functions and operations of the agency;
2. General goals and objectives, including outcome- related goals and objectives, for the major functions and operations of the agency;
3. A description of how the goals and objectives are to be achieved, including a description of the operational processes, skills and technology, and the human, capital, information, and other resources required to meet those goals and objectives;
4. A description of how the performance goals included in the plan required by section 1115(a) of title 31 shall be related to the general goals and objectives in the strategic plan;
5. An identification of those key factors external to the agency and beyond its control that could significantly affect the achievement of the general goals and objectives; and
6. A description of the **program evaluations** used in establishing or revising general goals and objectives, with a schedule for future program **evaluations.**" (Emphasis added)
7. (Chapter 3 of title 5, USC Sec. 306)

#### Annual Performance Plan

The requirement for annual performance plan strongly suggests the necessity for USAID's evaluations as a "means to be used to verify and validate measured values."

The GPRA secondly amended the law so the Director of the Office of Management and Budget (OMB) "shall require each agency to prepare an annual performance plan covering each program activity set forth in the budget of such agency. Such plan shall:

1. establish performance goals to define the level of performance to be achieved by a program activity;
2. express such goals in an objective, quantifiable, and measurable form unless authorized to be in an alternative form under subsection (b);
3. briefly describe the operational processes, skills and technology, and the human, capital, information, or other resources required to meet the performance goals;
4. establish performance indicators to be used in measuring or assessing the relevant outputs, service levels, and outcomes of each program activity;
5. provide a basis for comparing actual program results with the established performance goals; and
6. describe the means to be used to verify and validate measured values."

### Program Performance Report

State Department and USAID are required to produce an annual program performance report to “include the summary findings of those **program evaluations** completed during the fiscal year covered by the report.” (Emphasis added)

Thirdly, the GPRA requires “the head of each agency shall prepare and submit to the President and the Congress, a report on program performance for the previous fiscal year. Covering the actual results for the three preceding fiscal years, "Each report shall:

Review the success of achieving the performance goals of the fiscal year;

1. Evaluate the performance plan for the current fiscal year relative to the performance achieved toward the performance goals in the fiscal year covered by the report;
2. Explain and describe, where a performance goal has not been met (including when a program activity's performance is determined not to have met the criteria of a successful program activity . . . or a corresponding level of achievement if another alternative form is used)-
3. "(A) why the goal was not met;
4. "(B) those plans and schedules for achieving the established performance goal; and
5. "(C) if the performance goal is impractical or infeasible, why that is the case and what action is recommended;
6. Describe the use and assess the effectiveness in achieving performance goals of any waiver under section 9703 of this title; and
7. Include the summary findings of those **program evaluations** completed during the fiscal year covered by the report. (Emphasis added)

Additionally the GPRA passed into law a government-wide definition of the meaning of evaluation: “program evaluation' means an assessment, through objective measurement and systematic analysis, of the manner and extent to which Federal programs achieve intended objectives.”

### **Program Assessment Rating Tool (PART)**

Implementation of the GPRA by OMB places emphasis on evaluations that are *independent, of sufficient scope and quality, conducted regularly, and adequate to demonstrate strong evidence* of a program’s effectiveness.

The mission of United States Office of Budget and Planning is to assist the President in overseeing the preparation of the federal budget and to supervise its administration in Executive Branch agencies. In support of GPRA, OMB developed in 2002 a Program Assessment Rating Tool (PART). The PART is a questionnaire developed to examine

various factors that contribute to the effectiveness of a program and requires that conclusions be explained and substantiated with evidence.

The PART includes two questions specifically related to evaluations.

- Question 2.6 asks whether there “[a]re **independent evaluation** of sufficient scope and quality conducted on a regular basis or as needed to support program improvements and evaluate effectiveness and relevance to the problem, interest, or need.” (Emphasis added)

Question 4.5 asks if “**independent evaluations** of sufficient scope and quality indicate that the program is effective and achieving results.” (Emphasis added)

## **ANNEX 2: EVALUATION RESOURCES**

Over the past two decades, literature on evaluations of international development assistance has proliferated at an unprecedented scale. A number of books, monographs and reports have been published, which provide guidance for planning, conducting and utilizing evaluations of foreign assistance projects and programs. Managers planning new evaluations should consult them, when possible. They are usually available in major libraries or can be purchased through the Internet.

In addition, web-sites managed by bilateral and multi-lateral donor agencies, teaching and research institutions, and non-governmental organizations provide invaluable information on evaluations. Their main advantage is that the necessary information can be instantly accessed and downloaded. Some important web-sites on international evaluation are listed below:

<http://dec.usaid.gov>

This site maintained by USAID provides a goldmine of information on evaluations of foreign assistance. It contains three categories of documents, which are of particular interest to evaluators and their managers.

First, the site is a repository of most evaluations conducted by USAID since its beginning. Therefore it is a source of critical ideas, concepts, methodologies and evaluation modalities for planning and conducting evaluations. For example, an evaluation manager, who is preparing a SOW for an impact evaluation of a micro-enterprise project, can download earlier impact evaluations on the subject. Such evaluations will help him/her in identifying evaluation questions, determining the nature of the data and information needed and proposing a realistic time frame for evaluation. They will also provide him/her some idea of the resources needed for the proposed impact evaluation.

Second, the site also keeps a relatively large number of monographs, guidelines and reports on evaluation research and methodology issued by USAID. For example, it has several publications on innovative, low cost data collection methods that can be used to collect data and information in different evaluation settings. Various bureaus and offices of USAID have prepared guidelines on conducting evaluations, which are also available on this site.



Third, USAID has published 14 “TIPS” on different aspects of monitoring and evaluation. While they might not always contain the most recent information, the following TIPS might be particularly helpful:

- Role of Evaluation
- Building a Result Framework
- Rapid Appraisal Methods
- Direct Observation
- Group Interviews
- Key Informant Interviews
- Data Quality
- Participatory Evaluation
- Measuring Institutional Capacity
- Preparing a Scope of Work
- Evaluation Report

The above TIPS are also available on the site.

[www.oecd.org/dac/evaluationnetwork](http://www.oecd.org/dac/evaluationnetwork)

The Network on Development Evaluation of the donor Development Assistance Committee (DAC) promotes independent evaluation and is an important mechanism for exchanging ideas and information. In addition to containing DAC’s evaluation publications, this site provides links to (a) the evaluation departments of its member agencies, and (b) evaluation associations and societies in Africa, Americas, Asia and Europe. The site also lists resources for international evaluation.

[www.undp.org/eo](http://www.undp.org/eo)

This is the official site of United Nations Development Programme, which conducts evaluations of its projects and programs. It posts evaluations completed by UNDP on a wide range of development projects and programs. Moreover, it also posts UNDP’s publications on evaluation methodology.

[www.worldbank.org/oed](http://www.worldbank.org/oed)

The Independent Evaluation Group (IEG) is an independent unit within the World Bank which conducts rigorous and timely evaluations. In addition to providing general information about the activities of IEG, this site lists important publications of IEG, including on evaluation methodology. Its publications are excellent in terms of their substance and presentation and set models for evaluation research.

[www.worldbank.org/reference](http://www.worldbank.org/reference)

This site holds thousands of evaluations and analytical studies conducted by the World Bank in various countries. The quality of evaluations, particularly on economic development projects and programs, is quite good. In addition, the site also posts sectoral reviews, policy documents and other publications, which can be helpful to evaluators.

The World Bank has also published several books on evaluation, which can be purchased through its bookstore.

[www.dfid.gov.uk](http://www.dfid.gov.uk)

This site contains evaluations conducted by DFID and its contractors. It also provides access to a variety of working papers, policy studies and other studies undertaken under the auspices of the Department.

[www.acdi-cida.gc.ca](http://www.acdi-cida.gc.ca)

The evaluation division of the Canadian International Development Agency undertakes evaluations of its projects and programs. This site contains several evaluation guides issued by the division as well as its evaluation reports. However, the number of evaluations posted on the site is not as large as those to be found in the sites maintained by DFID or USAID. The site also has a series of reports entitled “What we’re learning”, which summarizes the key lessons on various subjects.

[www.3ieimpact.org](http://www.3ieimpact.org)

International Initiative for Impact evaluation is a new organization which promotes rigorous impact evaluations for making policy decisions. Although relatively recent in origin, it has also produced several working papers on impact evaluations, and plans to bring out a journal on evaluation research. It also plans to create a comprehensive data base on impact evaluations in near future. When operational, it will be of immense help to evaluators and evaluation managers in development agencies.

[www.gao.gov](http://www.gao.gov)

GAO has published several informative monographs on evaluation methodology that are available on this site. In addition, the site regularly posts evaluations undertaken by GAO. Although most of these evaluations focus on domestic subjects, some such as on education and health have relevance to development topics. GAO has also conducted evaluations of USAID programs.

[www.dipllopedia.state.gov](http://www.dipllopedia.state.gov)

This site contains information about evaluation activities of the State Department. The Department is now taking steps to provide increased access its staff on evaluations conducted by its various bureaus and offices.

[www.wmich.edu/evalctr](http://www.wmich.edu/evalctr)

Hosted by Western Michigan University, this site provides checklists for designing, budgeting, contracting, staffing, managing, and assessing evaluations of programs, personnel. It also provides information about collecting, analyzing, and reporting evaluation information. However, this site is more oriented towards domestic U.S. programs.

[www.fedeval.net](http://www.fedeval.net)

This site maintained by “Federal Evaluators,” an informal association of federal evaluation officials, provides information about the evaluation activities of federal government and lists resources that are available.

[www.eval.org](http://www.eval.org)

This site of the American Evaluation Association provides information about the multifaceted activities of the Association. It also lists evaluators who are available for contracting.

[www.usaid.gov/policy/ads/200](http://www.usaid.gov/policy/ads/200)

This site maintained by USAID provides links to the updated ADS 200 series. These series list USAID directives for performance management, monitoring and evaluation.