



PERFORMANCE MONITORING & EVALUATION

TIPS

INTRODUCTION TO EVALUATION AT USAID

ABOUT TIPS

These TIPS provide practical advice and suggestions to USAID managers on issues related to performance monitoring and evaluation. This publication is a supplemental reference to the Automated Directive System (ADS) Chapter 203.

INTRODUCTION

This TIPS will provide the reader with a general introduction to the purpose, role, and function of evaluation in the USAID program and project design and implementation cycle. It will provide background on why evaluation has become an important part of the effort to improve the effectiveness of foreign assistance programming. It will also provide links to other TIPS with more detailed guidance on when and why to evaluate, how to evaluate, uses of evaluation data, how to address common problems, and how to structure the evaluation's findings, conclusions, and recommendations.

WHAT IS EVALUATION?

USAID policy defines evaluation as follows:

“Evaluation is the systematic collection of information about the characteristics and outcomes of Assistance Objectives [policies, programs], projects, or activities in order to make judgments, improve effectiveness, and/or inform decisions about current and future programming.” (ADS 203.3.6)

To enhance USAID's capacity and commitment to evaluation, as well as the strength of the evidence evaluations produce, USAID emphasizes that an evaluation is an analytical effort that can be examined in a:

“systematic way to gain insights and reach conclusions about the development hypothesis (original design), utility of

performance monitoring (the value of the Performance Management Plan), factors in the development context that may have an impact on the achievement of results (explanation), and the types of actions that need to be taken to improve performance (future action).”¹ (ADS 203.3.8)

EVOLVING POLICY

In 2008 USAID placed a renewed emphasis on the need for “robust and objective evaluations that are fundamental to effective evidence-based and results-based programming” (USAID General Notice, November 10, 2008). This shift signaled the need for program and project evaluations to play a central role in managing for results

¹ Note: Comments in parentheses are added for emphasis.

and learning from experience. In order for evaluations to support this objective, they must be evidence-based and of high quality. Evaluation reports must be shared, discussed, analyzed, disseminated widely, and used to inform future programming and ultimately strengthen development outcomes that yield results.

This renewed emphasis on evaluation was further underscored by the Office of Management and Budget (OMB). In a memorandum released on October 7, 2009, the USG made a statement declaring its increased emphasis on *rigorous* independent program evaluations. OMB stated that, “evaluations should be conducted with sufficient rigor to determine whether government programs are achieving their intended outcomes as well as possible and at the lowest possible cost,” to “help policymakers and agency managers strengthen the design and operation of programs,” and to “help the Administration determine how to spend taxpayer dollars efficiently and effectively through investing more in what works and what does not.”

In summary, evaluation is a priority for USAID. The Agency is part of a broader, government-wide effort to incorporate more evaluation as an important tool for evidence-based decision-making. It is also a priority to share evaluations and learning across organizations in the international development community.

THE LINK BETWEEN EVALUATION AND PERFORMANCE MONITORING

There are important connections between monitoring and evaluation, with monitoring serving as a

stepping stone to evidence-based evaluation. Together, these complementary elements form an effective performance management system. While monitoring can tell managers what is happening (has revenue increased? are children immunized?), evaluation examines why. USAID has invested over two decades in building performance management systems that serve dual management and accountability purposes. Some have argued that if monitoring systems were truly effective, then evaluation would not be needed. USAID policies go a long way towards explaining that systems cannot rely solely on performance monitoring. A strong monitoring system will provide continuous feedback about progress towards results and specific targets for predefined indicators at every level of a program/project. While this is an invaluable management tool, it is frequently not enough to answer more challenging questions.

Foreign assistance programs are generally based on a set of assumptions and hypotheses (or causal if-then relationships) that reflect a theory of change. USAID’s performance management systems are developed based on this type of thinking, as is reflected in Results Frameworks (see TIPS 13 Building a Results Framework).

Evaluation is used for examining relationships between results levels, testing development hypotheses, and validating critical assumptions. Evaluation is in a better position than performance monitoring to challenge project design. It looks beyond planned outcomes to discover unanticipated effects, both positive and negative. Together, performance monitoring and evaluation can provide USAID with a comprehensive picture of the effectiveness of its interventions.

WHY EVALUATE? THE USES OF EVALUATION DATA

All USAID evaluations should generate evidence-based data that can be used by policy-makers, program and project managers, and other stakeholders to make critical programming and resource-allocation decisions. Some examples of why it is so important to conduct evaluations include the following:

- To assess the impact of a program or project on its intended beneficiaries. If the program or project did not have the intended effect, evaluation data can be analyzed to explain why and then be used to assist in the future design of similar programs or projects.
- To determine if a pilot intervention should or should not be replicated elsewhere. If, for example, evaluation findings show that a pilot project intervention is successful, USAID managers and their stakeholders may decide that it can be scaled up and replicated elsewhere in the country with adjustments made as necessary.
- To determine whether a project or set of related projects are appropriate and sufficient to reach an overall program Assistance Objective mid-way through the life of the program. Findings from mid-term evaluations can be used to adjust the overall strategy, redesign specific project activities, or to change the manner in which contributing projects are implemented.
- To understand the effectiveness of USAID-supported policy reforms undertaken by host governments to promote positive changes in the economy or key indicators of

sector improvement. Evaluations can be used to trace the effect of policy reform on regulations and laws, law enforcement, actual changes in behavior, and the subsequent improvement in the economy or targeted sector, ultimately leading to positive changes for intended beneficiaries. The findings from evaluation of policy-reform programs can be used to support the government case for reform or to make adjustments as needed.

- To help design follow-on projects or programs. Mid-term or final evaluations of an AO program or project yield valuable information that can help strengthen the design of the next stage of a program or project by isolating features that must be continued or curtailed, or new interventions that must be designed.
- To develop evidence-based lessons learned on a class of programs or projects. Multi-country evaluations of similar programs that have been implemented in several regions of the world can yield valuable data that can be used to generate solid lessons learned and assist USAID managers, and other decision-makers, in deciding whether or not it is valuable to continue funding such programs or projects.

In utilization-focused evaluations, an approach introduced by Michael Quinn Paton (1997), the timing, questions, design, methods, and presentation of evaluation data are all focused on the end-user to ensure maximum utilization of the data by policy makers, managers, and other key decision-makers who commission evaluations.

WHEN TO EVALUATE

The need for evaluation should be considered at an early stage in the design of programs and projects. USAID policies (ADS 203.3.3.1) describe a “complete” Performance Management Plan (PMP) as one that includes the identification of “possible evaluation efforts to complement performance monitoring.” Strengthening efforts to articulate and budget for evaluations in PMPs for programs and projects will help ensure that USAID is ready to meet this new evaluation requirement.

Consistent with its view of evaluation as a useful management tool, USAID policies (ADS 203.6.1) require program and project managers to undertake evaluations. They should be conducted whenever an evaluation would be the best way to answer questions needed to inform forward-looking decisions, examine causality, or arrive at summative judgments about the worth of a program/project, either “in progress,” or with the benefit of hindsight.

The decisions as to when to call for an evaluation and what questions to address are best made by those closest to the action. USAID staff and design teams on the ground know when they are testing an intervention to determine whether it should be scaled up. If so, they know whether or not it would benefit from building a comparison group into the projects plans. Likewise, USAID staff can also recognize when a project is not gaining traction on the problem it was designed to address. When it comes to the details of an evaluation such as determining the size of the scope, or identifying the need for an evaluation (i.e. pre-planned or *ad*

hoc), involved USAID staff often know best.

At the same time, USAID must ensure that it does not ignore the Government Performance and Results Act (GPRA), despite its many other responsibilities. This act mandates the use of evaluations to gain a broad empirical perspective on program and project effectiveness within a schedule consistent with the Agency’s mid-range and long-term planning cycles. USAID policies have therefore introduced a requirement for periodic evaluation at the program level:

“AO Teams must conduct at least one evaluation aimed at understanding progress or lack thereof and the types of actions that need to be taken to improve performance *during the life of each Assistance Objective (AO)*. (ADS 203.3.6.1)”

These policies also identify some key “triggers” for a program – or project – evaluation. These include situations where:

- A key management decision is required about the future direction of an activity, project, or program, but there is inadequate information to guide this decision-making process.
- Performance monitoring data indicates an unexpected result (positive or negative) that needs further explanation and/or elucidation.
- Feedback from partners and/or other stakeholders suggests that there are implementation problems, unintended consequences, or unmet needs.
- Issues of sustainability, cost-effectiveness, or relevance arise.
- The validity of the logframe, results frameworks, development

hypotheses, or critical assumptions is questioned.

- Periodic Portfolio Reviews have identified key questions that need to be answered.
- Extracting lessons is important for other operating units or for future programming.

In summary, evaluations can be conducted whenever justified and whenever a management decision and/or design process can be assisted by its results.

There is also a new, real, and timely challenge facing Agency staff. New policies emphasize planning, conducting, and utilizing one, or a cluster of, evaluations focused on the AO level rather than on individual projects. Focusing on the AO level implies addressing AO level progress, the validity of development hypotheses at this level, unanticipated results and/or cost effectiveness, redundancy, and synergy across a program.

INTEGRATING EVALUATION PLANNING INTO PROGRAM AND PROJECT DESIGN

Program level evaluation, as well as stronger project level evaluation, will benefit from the kind of thoughtful planning and teamwork that characterize the best of USAID experiences (for example in developing strategies, AOs, and designing implementation mechanisms). During the design phase, managers should consider the following:

- How to fund program/project evaluations.
- Timing.
- Evaluation team composition, with a focus on identifying team members with complementary expertise. USAID intends to draw more heavily on internal staff to identify, design, participate in, and learn from evaluations.

- If and when comparisons over time or between target and non-target groups/areas are warranted (see TIPS 19 Impact Evaluation).
- How the collection of data required for assessing outcomes in evaluations can be integrated into program and project performance monitoring.

In this regard, it is important to focus on when and how baseline and performance data are collected and analyzed across several projects, particularly in reference to the host government, beneficiaries, and other stakeholders.

For example, decisions on timing may need to address multiple needs. One option is to allow for some flexibility to undertake evaluations on an ad hoc basis when circumstances indicate that managers would benefit from this type of empirical and analytic input.

At the same time, Missions and Offices may want to make more deliberate plans for evaluating all or segments of a program at a point where doing so will help inform a “next round” of program and activity designs, the development of associated RFPs (requests for proposals), RFAs (requests for application), or other solicitation devices. In practice, this may mean scheduling evaluation activities four to six months before a new strategic plan is required. Coordinating the timing of project and program level evaluations to inform updated program strategies may mean that some evaluations will be undertaken midway through projects. In other cases, managers may have the luxury of letting some time pass between the end of a project and its evaluation. This may increase the likelihood of learning what elements and benefits have been sustained beyond the point where USAID funding ended.

During the planning process, managers have an opportunity to consider what role USAID’s own staff and other key stakeholders will play in such evaluations. USAID staff involvement, in particular, helps to ensure that evaluation findings are fully utilized in planning and implementing “next round” strategies.

USAID managers also have the option of using external evaluation teams. This option is appropriate when decision-makers believe that a program or project would benefit from an independent analysis, including situations where USAID may have reason to believe that an evaluation will raise issues that it would rather have raised by a third party.

Joint or collaborative evaluations are another increasingly attractive option given commitments to the Paris Declaration. Donors and host countries alike are encouraged to transparently collaborate on evaluations. USAID policies provide ample space for this type of evaluation, and guidance provided by the OECD/DAC on the conduct of such evaluations provides helpful support in this area (see <http://www.oecd.org/dataoecd/28/14/37484787.pdf>).

TYPES OF EVALUATION

There are many types of evaluations, (formative, summative, impact) as well as styles of evaluation (participatory, collaborative, utilization focused, etc.) that are conducted. While it is not possible to provide a comprehensive taxonomy of the various types of evaluation, the text box² (entitled Selected Evaluation Types and Concepts) below provides

² Based on Rossi, Peter H., et al, Evaluation: A Systematic Approach, Sixth Edition, Sage Publications, Thousand Oaks, CA, pg.36.

definitions for some common evaluation types and concepts.

Mid-term evaluations are conducted mid-way during the life cycle of a program or project. They can include process questions that are usually conducted as part of a **formative evaluation**, such as an assessment of implementation issues and partner cooperation (see textbox definition of formative evaluations), as well as a determination of why indicator target values are not being reached and the likelihood of reaching program or project objectives. Information from mid-term evaluations can be used to strengthen the interventions for the remainder of the program or project.

Final or summative evaluations are conducted at the end to determine whether, how, and why program or project outcomes have or have not been met. There is a tension in USAID between whether or not to conduct a mid-term or final evaluation and when to conduct such evaluations, based on the frequently-occurring need to have early information with which to design follow-on programs or projects. Because of the length of time it takes to design and complete a new project, USAID managers are tempted for reasons of cost and time to conduct one evaluation serving multiple needs; such a decision is best left to the USAID program or project manager.

Impact evaluations are usually conducted at a point after the complete program or project outcomes on targeted beneficiaries are expected to be fully realized. The timing depends on the development hypothesis, the nature and timing of when these effects should be realized, and the context in which the program or project is implemented.

SELECTED EVALUATION TYPES AND CONCEPTS

Formative Evaluation: Evaluative activities undertaken to furnish information that will guide program improvement. Initiated before implementation begins or near the start of a project or program, they are used to refine the intervention and help guide the approach and initial implementation. A formative evaluation is a specific type of process evaluation.

Process Evaluation: An evaluative study that answers questions about program operations, implementation, and service delivery. Also known as an implementation assessment.

Summative Evaluation: Evaluative activities undertaken to render a summary judgment on certain critical aspects of the program's performance: for instance, to determine if specific goals and objectives were met. Summative evaluations are undertaken in the context of final evaluations.

Impact Evaluation: Evaluative activities undertaken to answer questions about program outcomes and their impact on changes in conditions (e.g., for people and/or in the prevailing social and/or, economic, environment, political conditions) targeted by the program. Increasingly, impact evaluations are using more rigorous evaluation methods that use experimental or quasi-experimental methods to establish the counterfactual (See TIPS # 19: Impact Evaluation).

Collaborative Evaluation: An evaluation that is organized and conducted to include multiple parties in design and or implementation. For example, today's collaborative evaluations are designed and executed by donors involved in supporting one overall program through its donor coordination efforts.

Participatory Evaluation: An evaluation organized as a team project in which the evaluator and representatives of one or more stakeholder groups work collaboratively in developing the evaluation plan, conducting the evaluation, or disseminating and using the results (See TIPS #1: Participatory Evaluation).

Empowerment Evaluation: A participatory evaluation in which the evaluator's role specifically includes consultation and facilitation directed toward the development of the participating stakeholders to conduct evaluation on their own, to use it effectively for advocacy and change, and to have some influence on a program that affects their lives.

Utilization-focused Evaluation: A concept and approach that focuses on the end-users of the evaluation and their intended use of findings and conclusions. Employing this focus determines the timing of when the evaluation is conducted based on end-user needs, evaluation design and methods, and how data, findings and conclusions are presented to facilitate decision-making.

The type of evaluation that is most appropriate depends on the purpose, the resources allocated to the task, the timing, and who is expected to actually make use of the evaluation findings, conclusions and recommendations. It is nevertheless useful to understand the dimensions along which types of evaluations can be arrayed. In simple terms, these dimensions focus on **when, what and how we evaluate**.

WHEN

As discussed above, the "when" dimension is invariably linked with an evaluation's purpose. The terms formative and summative are most often used in this regard. Purely formative evaluations are undertaken during program implementation with the aim of improving on-going efforts. Evaluations that blend formative and summative are often used to create

a bridge from one program or project to the next, or between phases. They serve the dual objectives of increasing accountability and improving learning. Purely summative evaluations tend to be undertaken towards the end of a program or project, or after USAID funding has ended.

WHAT

The “what” dimension ranges on a continuum from process to outcome. This continuum is often reflected in the **type** of evaluation questions included in a Statement of Work. (See TIPS No. 3: Preparing an Evaluation Statement of Work).

Evaluation of Process Questions

Formative evaluations are used at the beginning of a program or project to address process questions about how well program or project implementation and operations are going with respect to their contribution to positive outcomes. Generally speaking, process questions are of greatest interest during implementation, when procedures can be changed and relationships improved. At the same time, there can be merit in asking process questions retrospectively, particularly when the purpose of an evaluation is to learn in order to improve the design of a class of programs or projects across a number of districts, regions, or countries.

Evaluation of Outcome Questions

At the other end of this continuum are outcomes (as seen in project Logframes and in Results Frameworks as AOs and IRs). These include both planned and unintended consequences of programs or specific projects. Planned outcomes are often linked by a sequence of “if-then”

hypotheses. Most results frameworks include at least two to three outcomes in a hierarchical chain, and an evaluation may wish to explore several of those outcomes as well as the degree to which USAID’s intervention versus other factors play a causal role. Outcomes occur at different levels and types of interventions, including programs, policies, and projects.

Summative evaluations are usually conducted at the conclusion of either a program, policy, or project intervention to determine if outcomes have been achieved. They also provide important data to promote development learning that should be shared.

Outcome is a general term. In the USAID context, both Intermediate Results (IRs) and Assistance Objectives (AOs) are intended outcomes of one or more activities that comprise an overall program. The outcome of a specific project that is part of a larger program is often expressed as an IR outcome. See TIPS 13: Building a Results Framework for further definition and explanation.

Host country partners and other donors may use terms such as goals or development objectives to connote these kinds of results. In addition, some in the development community use the term “impact” to refer to long-term outcomes or to outcomes that directly affect target populations.

In summary, evaluations can be conducted to focus on questions anywhere along this process-outcome continuum by focusing exclusively on process or outcome, or blending the two. “What” you are evaluating, either process or outcome questions, determines the type of evaluation that should be conducted and “how” the evaluation design is structured.

HOW

A final dimension (particularly relevant for scaling evaluations) focuses on “how” the evaluation is structured with respect to answering questions as to whether change occurred, and, if so, what causal factors brought about that change. The types of evaluation questions outlined determine the how an evaluation is structured.

In practice, existing performance monitoring systems may or may not track “process” factors that an evaluation needs to address. These factors may include the tracking of “outcomes” at all levels of a results framework, or the gathering of baseline data before (rather than at some point after) USAID’s program or project activities (i.e., when intervention or treatment) begins. Where gaps exist, evaluators have to make choices about how they will determine whether the outcomes of interest in a program or project show evidence of change.

For example, there are situations in which it is too late to begin collecting baseline data since the intervention had already begun, or where data were collected too late to be considered unaffected by the intervention. In those cases, “how” decisions involve choosing among existing data series. This data must go back far enough to register a change, if it occurred, or must be able to reconstruct the pre-program or pre-project baseline situation. This can be done through reconstruction techniques such as the use of recall, among other approaches.

Being able to prove whether change occurred is often essential for an evaluation, but proof of change is not proof of causality. At best, it *correlates* the implementation of a USAID program or project with an outcome, i.e., it says they happened in tandem.

A second aspect of the “how” dimension focuses on whether and to what extent USAID’s intervention caused specific outcomes. Demonstrating causality means that the evaluator must show that in the absence of USAID’s intervention the change would not have occurred. Along the continuum of evaluations that do this are studies involving rigorous comparisons between USAID target groups or areas and non-recipient groups or areas (see TIPS No. 19: Impact Evaluations).

On the continuum just short of impact evaluations are a range of evaluation designs that utilize time-series data and econometric techniques to determine the impact of various influences, including USAID’s intervention, on an outcome over time. Still other options are evaluation designs that seek to rule out alternative causes by other means, including methods that are forensic in nature. They start from the fact that proof of changes in an outcome exist and work backwards through a range of alternative possible causes to determine which had the strongest influence or how they collectively contributed to that result.

In instances where USAID works in concert with other donors on a particular large-scale or sector-wide program, it is more difficult to single out the causality of USAID-specific interventions alone in terms of whether or not (or how) final program outcomes have been achieved. In these instances, it is sometimes more appropriate to look for attribution of USAID’s specific interventions in relation to overall program outcomes. A collaborative evaluation design among donors is a useful way to approach such issues.

ENSURING HIGH QUALITY

A high-quality evaluation is one that employs the optimal research design to answer specific evaluation questions, appropriate use of data collection methods, rigorous analytical techniques, and the careful use of evidence to develop findings, conclusions, and recommendations. The qualifications and experience of evaluation team members are also critical. This TIPS does not attempt to cover all of the many steps evaluators must take to produce a high-quality evaluation. Readers are referred to the TIPS series on different evaluation topics and the publications in the suggested reading section for more information on designing and executing high quality evaluations.

There are a number of common problems that evaluators may face during the course of conducting the evaluation which can affect evaluation quality. These include lack of baseline data from the onset of implementation, inadequate sample sizes from which to determine impact, too little time accorded to carry out an evaluation, problems maintaining a valid comparison group over time, and so on. Table I briefly reviews some of the most common problems and some means to address them.

ETHICAL ISSUES

Evaluators and evaluation users are sometimes confronted with ethical issues that, if not effectively addressed, can affect the validity or credibility of the evaluation and even cause some harm. Evaluation can be seen as a political process: all stakeholders have an interest in being seen in a positive light while at the same time benefiting from the insights and knowledge produced by a well-done evaluation. There is

substantial room for a variety of ethical conflicts of interests to arise. A few are highlighted here, but it is a good idea to think through, and address the potential for such conflicts in evaluations. Some are:

- Failure to gain informed consent from respondents to a survey.
- A cognizant officer’s technical representative (COTR) intervenes or rewrites sections of the report to put the project in a more favorable light.
- Opinions and views gained with the promise of confidentiality are reported in a manner that identifies the respondent.
- The evaluation team publicly alleges waste, fraud, and abuse on the part of implementers or host government organizations.
- Findings and analysis are based on “cooked” data, rather than actual data collected or found and analyzed during the evaluation.

The best way to deal with ethical issues is to take steps in **advance** to prevent them from occurring. Professional evaluators should be required to adhere to codes of conduct established by the American Evaluation Association (AEA). These include the Program Evaluation Standards (1994) and the AEA Guiding Principles for Evaluators (1995, 2004). These standards and guides can also help to identify means to address ethical issues as they arise during the course of the evaluation. Additionally, before the evaluation team begins, a full discussion should

TRANSPARENCY

USAID policy is to openly share and discuss evaluation results with relevant partners, as well as other donors and stakeholders (unless there are unusual and compelling reasons not to do so).

be held on roles, responsibilities, and expectations among evaluators and evaluation stakeholders, based on the scope of work, to help prevent problems. Lastly, everyone involved should understand that the independence and integrity of the evaluation process is essential to the credibility and validity of the evaluation report.

COMMUNICATING AND SHARING FINDINGS TO ENSURE LEARNING

Communicating evaluation results effectively is critical if they are to be used. Evaluators and USAID staff that commission evaluations need to be proactive in seeking out opportunities to interject evaluation results into relevant management discussions and decisions. They also need to be creative in tailoring a communication strategy to fit the diverse audiences' needs.

When formal evaluation reports are prepared, they should be succinct, appealing, readily understood, and useful for decision-makers (Refer to TIPS No.17: Constructing an Evaluation Report for guidance). Along with narrative, graphics and other types of illustrations should be used to convey key points.

Evidence-based findings, conclusions, and recommendations must be clearly identified and distinguished from each other. Making these distinctions enables readers to trace the reasoning used by the evaluators in reaching conclusions and proposing recommendations, thus

making the evaluation more transparent.

Share evaluation results widely. USAID policy is to openly share and discuss evaluation results with relevant partners as well as other donors and stakeholders (unless there are unusual and compelling reasons not to do so). Such transparency enables others to learn and benefit from the evaluation's results and facilitates their broader use. Evaluation reports should be translated into the language of key counterparts.

Use oral briefings. Briefings are almost always more effective than written reports for presenting evaluation results, and their use is suggested whenever possible. By creating a forum for discussion among relevant actors, briefings create momentum for action. Most importantly, briefings fit the way busy managers normally operate; they rarely have time to sit and read lengthy documents and moreover are used to making decisions jointly with others in meetings.

Use multiple communication techniques. Using written reports and briefings to communicate evaluation results is commonplace, but also consider using less traditional techniques that may be effective at feeding evaluation findings into ongoing decision-making or that aim at sharing evaluation results more broadly. For example, consider using websites, senior managers' bulletins, memoranda, email Listserve messages, question-and-answer statements, press releases, op-ed items in newspapers, speeches, written testimony, newsletters,

articles in professional journals, brown-bag lunches, conference presentations, videotapes, or computerized evaluation presentations.

Submit evaluation reports to the DEC. The automated development experience clearinghouse (DEC) – which includes thousands of evaluation reports – is a vital aspect of the Agency's capacity to learn and share experiences across operating units and with the broader development community. Operating units are required to submit in electronic form all evaluation reports, executive summaries of evaluations, other documents prepared at the conclusion of an evaluation activity, operating unit's (or counterpart agency's) response to evaluation reports, and action decisions arising from evaluation activities.

REVIEW AND USE OF EVALUATION RESULTS

Operating units have the primary responsibility for responding to and using an evaluation, including:

- Systematically reviewing the key findings, conclusions, and recommendations.
- Identifying which are accepted and supported and which are not.
- Identifying specific management actions and assigning clear responsibilities for undertaking them.
- Determining whether any revisions are necessary in strategy or activities.

TABLE I. COMMON EVALUATION PROBLEMS

Problem	Description	Ways to Address
<i>Too many questions, so little time</i>	Evaluation Statements of Work often contain many complex questions, but the funding and time is often limited to two to three weeks in the field with a correspondingly short time for data analysis, drafts, reviews, and the final report. A survey of evaluators showed that the single most important constraint to doing high quality evaluation was the failure of the issuing client to allocate sufficient time and budget for the task.	Statements of Work must provide a realistic and adequate budget for both time and resources based on the nature and scope of the evaluation purpose and questions. TIPS # 3: Preparing an Evaluation Statement of Work provides guidance.
<i>Inadequate baseline data</i>	Lack of or inadequate baseline data against which to measure changes in the target population is probably the most common problem faced by evaluators and one of the most serious threats to the validity of the evaluation. Baseline data by itself is not sufficient to assess attribution, but without it, the evaluator cannot measure change in any rigorous way	Reconstructing baseline data can be done by using secondary data, individual recall, participatory group techniques to reconstruct history and assess changes produced by the intervention, and key informant interviews. Data from any one method must be used cautiously. Evaluators should triangulate the estimates of reported information by using multiple data sources to increase the validity of the reconstructed baseline. TIPS # 5: Rapid Appraisal defines and discusses triangulation methods.
<i>Dangerous program settings prevent access to collecting evaluation data</i>	Many of the USAID's largest assistance programs are in countries that are unstable or racked with internal conflict. Reaching key segments of the population to collect data may be dangerous and, even if possible, citizens may be afraid to speak to outsiders.	In such instances, evaluators must work with stakeholders to discuss alternative data sources and data collection methods that are reasonable and acceptable under such conditions. A special TIPS will be written on this subject.
<i>Maintaining comparison group differences</i>	Effective use of comparison groups in impact evaluations requires both stability in the task environment and careful management from beginning to end, often over a 3 to 5 year period. If the project or program is providing desirable benefits, it is difficult to prevent individuals in comparison groups from securing those benefits. In other situations, program effects in the target group may spillover to the control group selected for a comparison. This results in underestimation of program impact since the control group will appear better-off than they would have.	In some cases, "spillovers" can be mapped and measured, and then taken into account during the analysis of data from the target and control groups. However, the most effective means to deal with such an issue is to control it in advance through an evaluation design that selects treatment and control groups that are unlikely to significantly interact with one another. See TIPS # 19: Impact Evaluation.
<i>Disagreements on findings, interpretation and conclusions</i>	Serious disagreements between stakeholders and the evaluation team on findings or interpretation/analysis and conclusions can threaten the credibility and usefulness of evaluations.	Hold a facilitated discussion on the relationship of the data (evidence) and its analysis and interpretation to the findings, and how these formed the basis of the conclusions. The usual practice is for the report to identify those points of disagreement in a foot note or annex.

Recommended Reading

Michael Bamberger, Jim Rugh, and Linda Mabry. *Real World Evaluation: Working Under Budget, Time, Data, and Political Constraints*. Sage Publications, Inc., Thousand Oaks, CA. 2006.

Eleanor Chelimsky and William R. Shadish (eds.). *Evaluation for the 21st Century: A Handbook*. Sage Publications, Inc., Thousand Oaks, CA. 1997.

Michael Quinn Patton. *Utilization-Focused Evaluation*. 3rd ed. Sage Publications, Inc. Thousand Oaks, CA. 1997.

Peter H. Rossi, Howard E. Freeman, Mark W. Lipsey. *Evaluation: A Systematic Approach*. 6th Edition. Sage Publications, Inc., Thousand Oaks, CA. 1999.

For more information:

TIPS publications are available online at [insert website].

Acknowledgements:

Our thanks to those whose experience and insights helped shape this publication including USAID's Office of Management Policy, Budget and Performance (MPBP). This publication was authored by Richard Blue, Ph.D., Patricia Vondal, Ph.D., and Molly Hageboeck, of Management Systems International.

Comments can be directed to:

Gerald Britan, Ph.D.

Tel: (202) 712-1158

gbritan@usaid.gov

Contracted under RAN-M-00-04-00049-A-FY0S-84

Integrated Managing for Results II