USAID EVALUATION POLICY

Bureau for Policy, Planning, and Learning

PREFACE

Built on our agency's rich tradition of evaluation, this policy sets out an ambitious recommitment to learn as we "do," updating our standards and practices to address contemporary needs. In an increasingly complex operating environment, the discipline of development demands a strong practice and use of evaluation as a crucial tool to inform our global development efforts, and to enable us to make hard choices based on the best available evidence.

This policy was developed in response to calls from within USAID. The development professionals who apply their best thinking to solve hard problems know that we can learn more systematically from our work, and that we can more rigorously and credibly document our programs' effectiveness. I have been inspired to see the broad and active engagement throughout the organization in establishing high standards that reflect an enduring commitment to using well the resources entrusted to us.

I have great expectations for the work of USAID. I expect us to succeed in some of our efforts, and to fall short in others. I expect a strong evaluation function and feedback loop that enables us to be accountable in both cases, and to learn from each so that we can make continuous improvements. We can do this only with evidence and data to inform our decisions, and with unprecedented transparency about what we have learned and where.

That is why I am so excited to share this policy. In it you will find more demanding evaluation requirements, ensuring that the majority of our program resources are subject to evaluation. You'll learn of our commitment to high methodological standards that are clear from the design stage, and that ensure to the extent possible that a different evaluator using the same methods would arrive at similar findings and conclusions. We will be unbiased, requiring that evaluation teams be led by outside experts and that no implementing partner be solely responsible for evaluating its own activities. We will be transparent, registering all evaluations and disclosing findings as widely as possible, with standard summaries available on the website in a searchable form. To support these new standards, we will reinvigorate our training, ability to access technical expertise, and investment in evaluation.

Importantly, it is our hope that you will find this policy in and of itself a basis for our own organizational learning. We will continue to improve and make adjustments as we implement so that we can incorporate new ideas and maintain its relevance for USAID.

In the end, the measure of our success will not be predicated on the number of evaluations done, or stored within a database, or even solely upon the quality of the findings. We'll be successful if and when the evaluation work of USAID contributes to greater development effectiveness. When implemented, our vision of success is that this policy will make us better able to identify areas where we are more capable of achieving development results, and those where we have less success and therefore need to improve. Our success will depend on our ability to use evaluation findings to strengthen our efforts and sharpen our decision-making. With the implementation of this policy, we expect a step change in the quantity and quality of evaluation findings that inform our own strategies, program design, and resource allocation decisions; and we will contribute to the global community with new, practical and rigorous knowledge.

I feel very privileged to be a part of this team and part of this commitment.

– Administrator Shah

ACKNOWLEDGMENTS

1. CONTEXT

USAID stewards public resources to promote sustainable development countries around the world. Reflecting the intent of the authorizing legislation of the U.S. Agency for International Development (the Foreign Assistance Act of 1961, as amended) and embodying the aims of the current National Security Strategy, the Presidential Policy Directive on Global Development, and the Quadrennial Diplomacy and Development Review USAID pursues this goal through effective partnerships across the U.S. Government, with partner governments and civil society organizations, and with the broader community of donor and technical agencies. The Agency applies the Paris Declaration principles of ownership, alignment, harmonization, managing for results and mutual accountability.

To fulfill its responsibilities, USAID bases policy and investment decisions on the best available empirical evidence, and uses the opportunities afforded by project implementation to generate new knowledge for the wider community. Moreover, USAID commits to measuring and documenting project achievements and shortcomings so that the Agency's multiple stakeholders gain an understanding of the return on investment in development activities.

USAID recognizes that evaluation, defined in Box 1, is the means through which it can obtain systematic, meaningful feedback about the successes and shortcomings of its endeavors. Evaluation provides the information and analysis that prevents mistakes from being repeated, and that increases the chance that future investments will yield even more benefits than past investments. While it must be embedded within a context that permits evidence-based decision making, and rewards learning and candor more than superficial success stories, the practice of evaluation is fundamental to the Agency's future strength.

This policy builds on the Agency's long and innovative history of evaluation, and seeks to redress the decline in the quantity and quality of evaluation practice within the Agency in the recent past. The number of evaluations submitted to USAID's Development Experience Clearinghouse (DEC) decreased from nearly 500 in 1994 to approximately 170 in 2009, despite an almost three-fold increase in program dollars managed. Over that period, the Agency's evaluation activities have been subject to both internal and external critiques regarding methodological quality, objectivity, access to evaluation findings, and use of evaluation recommendations for decision making.

This policy responds to today's needs. High expectations exist for respectful relationships among donors, partner governments, and beneficiaries. Many stakeholders are demanding greater transparency in decision-making and disclosure of information. Development activities encompass not only the traditional long-term investments in development through the creation of infrastructure, public sector capacity, and human capital, but also shorter-term interventions to support and reinforce stabilization in environments facing complex threats. All of these features of the current context inform a policy that establishes higher standards for evaluation practice, while recognizing the need for a diverse set of approaches.

This policy is intended to provide clarity to USAID staff, partners and stakeholders about the purposes of evaluation, the types of evaluations that are required and recommended, and the approach for conducting, disseminating and using evaluations. Intended primarily to guide staff decisions regarding the practice of evaluation within projects managed by USAID, it also serves to communicate to implementing partners and key stakeholders a new approach to evaluation.

This policy works in concert with existing and pending Agency policies, strategies and operational guidance, including those regarding project design, evaluation-related competencies of staff, performance monitoring, knowledge management, and research management. It draws in significant ways on the evaluation principles and guidance developed by the OECD/DAC Evaluation Network. In addition, the policy is consistent with the Department of State Evaluation Policy, and USAID will work collaboratively with the Department of State Bureau of Resource Management to ensure that the organizations'

guidelines and procedures with respect to evaluation are mutually reinforcing. USAID also will work closely with the Department of State's Office of the Director of U.S. Foreign Assistance in its efforts to strengthen and support sound evaluation policies, procedures, standards and practices for evaluation of foreign assistance programs.

Box 1: Concepts and Consistent Terminology

To ensure consistency in the use of key concepts, the terms and classifications highlighted below will be used by USAID staff and those engaged in USAID evaluations.

Evaluation is the systematic collection and analysis of information about the characteristics and outcomes of programs and projects as a basis for judgments, to improve effectiveness, and/or inform decisions about current and future programming. Evaluation is distinct from assessment, which may be designed to examine country or sector context to inform project design, or an informal review of projects.

- *Impact evaluations* measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a "treatment" or a "control" group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.

- *Performance evaluations* focus on descriptive and normative questions: what a particular project or program has achieved (either at an intermediate point in execution or at the conclusion of an implementation period); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to program design, management and operational decision making. Performance evaluations often incorporate before-after comparisons, but generally lack a rigorously defined counterfactual.

- *Performance monitoring* of changes in performance indicators reveals whether desired results are occurring and whether implementation is on track. In general, the results measured are the direct and near-term consequences of project activities.

- *Performance indicators* measure a particular characteristic or dimension of project results (outputs or outcomes) based on a project's results framework and underlying theory of change. In general, outputs are directly attributable to the program activities, while project outcomes represent results to which a given program contributes but for which it is not solely responsible.

- *Performance management* ("Managing for Results") is the systematic process of monitoring the achievements of program activities; collecting and analyzing performance information to track progress toward planned results; using performance information and evaluations to influence decision-making and resource allocation; and communicating results to advance organizational learning and communicate results to stakeholders.

NOTE: In referring to "projects" throughout the document, the term is used to mean a set of planned and then executed interventions identified through a design process, which are together intended to achieve a defined development result, generally by solving an associated problem or challenge. The term "project" does not refer only or primarily to an implementing mechanism, such as a contract or grant.

2. PURPOSES OF EVALUATION

Evaluation in USAID has two primary purposes: accountability to stakeholders and learning to improve effectiveness.

Accountability: Measuring project effectiveness, relevance and efficiency, disclosing those findings to stakeholders, and using evaluation findings to inform resource allocation and other decisions is a core responsibility of a publicly financed entity. For evaluation to serve the aim of accountability, metrics should be matched to meaningful outputs and outcomes that are under the control or sphere of influence of the Agency. Accountability also requires comparing performance to ex ante commitments and targets, using methods that obtain internal validity of measurement, ensuring credibility of analysis, and disclosing findings to a broad range of stakeholders, including the American public.

Learning: Evaluations of projects that are well designed and executed can systematically generate knowledge about the magnitude and determinants of project performance, permitting those who design and implement projects, and who develop programs and strategies – including USAID staff, host governments and a wide range of partners – to refine designs and introduce improvements into future efforts. Learning requires careful selection of evaluation questions to test fundamental assumptions underlying project designs, methods that generate findings that are internally and externally valid (including clustering evaluations around priority thematic questions), and systems to share findings widely and facilitate integration of the evaluation conclusions and recommendations into decision-making.

These two purposes can be achieved simultaneously and span all projects. However, neither of these purposes can be achieved solely through the evaluation function. Each requires intentional actions by senior management to foster a culture of accountability and learning, and to provide appropriate incentives (and minimize disincentives) for staff at all levels.

3. BASIC ORGANIZATIONAL ROLES AND RESPONSIBILITIES

Each of the Agency's operating units that implement development projects will comply with this policy, supported by a set of central functions. Operating units will:

- Identify an evaluation point of contact. This individual will be responsible for ensuring compliance with the policy across the breadth of the operating unit's projects, and will interact with the regional and technical bureau points of contact and the Bureau of Policy, Planning and Learning, Office of Learning, Evaluation and Research (PPL/LER). The time allocated to this function should be commensurate with the size of the evaluation portfolio being managed.
- Invest in training of key staff in evaluation management and methods through Agency courses and/or external opportunities.
- Actively encourage staff to participate in relevant evaluation communities of practice for knowledge exchange.
- Develop, as needed, the guidance, tools and contractual mechanisms to access technical support specific to the types of evaluations required for the country, region or topical area in the domain of the operating unit. In general, this will require collaboration between the Program and Technical Offices. USAID missions will prepare a Mission Order on evaluation describing the context-specific approaches and expectations regarding evaluation.
- Prepare on a yearly basis an inventory of evaluations to be undertaken during the following fiscal year, as well as those completed. In general, the evaluations will be identified through the preparation of Performance Management Plans. The information will be included in the annual Performance Plan and

Report (PPR). Performance Plan and Report guidance will indicate the specific information to be supplied.

- Develop, through the Program Office (as defined in ADS 100), a budget estimate for the evaluations to be undertaken during the following fiscal year. On average, at least 3 percent of the program budget managed by an operating unit should be dedicated to external evaluation.[1]

- Ensure that final scopes of work for external evaluations adhere to the standards described below (See Section 4). In general, this will require collaboration between the Program and Technical Offices. The Program Office will engage the regional and technical bureaus in reviews of evaluation scopes of work. Except in unusual circumstances, as determined by the leadership of the operating unit, the Program Office will manage the contract or grant relationship with the evaluation team or consultant.

- Ensure, through the Program Office, that evaluation draft reports are assessed for quality by management and through an in-house peer technical review, and that comments are provided to the evaluation teams.

- Ensure, through the Program Office, that evaluation final reports (or reports submitted by evaluators to USAID as their final drafts) and their summaries are submitted within three months of completion to the DEC.

- Ensure, through the Program Office, that evaluation data are warehoused for future use.

- Integrate evaluation findings into decision making about strategies, program priorities, and project design. In general, the Program Office will take responsibility for this function.

- Make available all relevant information for technical audits of evaluation practices, as described below.

- Participate in the Agency-wide process of developing an evaluation agenda.

Each of the technical and regional bureaus will:

- Identify an evaluation point of contact. This individual will be responsible for ensuring compliance with the policy across the breadth of the operating unit's projects, and will interact with the PPL/LER. The time allocated to this function should be commensurate with the size of the evaluation portfolio being managed.

- Invest in training of key staff in evaluation management and methods through Agency courses and/or external opportunities.

- Participate in an evaluation community of practice for knowledge exchange.

- Organize, on request of the mission Program Offices, reviews of evaluation scopes of work and draft evaluation reports.

- Participate in the Agency-wide process of developing an evaluation agenda.

PPL/LER is an institutional source of guidance, support and quality assurance for the design, conduct, dissemination, and synthesis of evaluations. PPL/LER will:

- Develop training curricula and evaluation tools that have wide application across the Agency's portfolio. Identify opportunities for external training in specialized topics.

- Organize and lead the Evaluation Interest Group and other cross-Agency evaluation-related knowledge networks.

- Develop and/or update, with the Office of Human Resources, capabilities statements for evaluation specialists and senior evaluation specialists.

---

[1] An external evaluation is one that is commissioned by USAID, rather than by the implementing partner, and in which the team leader is an independent expert from outside of the Agency, who has no fiduciary relationship with the implementing partner.

- Organize technical resources for evaluation that can be accessed through a flexible mechanism. This includes, among other services: developing appropriate technical specifications for competitively procured evaluation expertise, reviewing and approving evaluation scopes of work, coordinating access to evaluation services, and providing estimates of evaluation costs.
- Respond on a priority basis with technical input for evaluation design and implementation, particularly for Presidential Initiatives and large country programs. This includes providing input into the requests for proposals for mechanisms to access technical support for evaluations.
- At any time, and particularly when requested by the Administrator, undertake or require a performance and/or impact evaluation of any project within the USAID portfolio.
- Undertake occasional thematic or meta-evaluations, to generate recommendations regarding Agency priorities, policies and practices. These evaluations will adhere to the standards described below.
- Undertake occasional post-implementation evaluations, to examine long term effects of projects.
- Organize occasional external technical audits of operating units' compliance with the evaluation policy, including through random checks of the technical quality of evaluation scopes of work, evaluation reports and utilization of evaluation findings.
- Provide clearance on exceptions to the requirement of public disclosure of evaluation findings, in cases where national security considerations and/or proprietary information is involved.
- Lead the preparation on a biannual basis of an Agency-wide evaluation agenda. Broad input from across the Agency, and from external stakeholders, will be sought during this process.
- Prepare an annual report for the Administrator highlighting recent key evaluation practices and findings, and changes and challenges in evaluation practice. Information for this will come from the PPR, among other sources.
- Serve as the main point of contact on evaluation with the Department of State, interagency partners, the Office of Management and Budget, the OECD/DAC Evaluation Network and the International Initiative for Impact Evaluation.
- Participate with F in the conduct of whole of government evaluations and in joint cross-cutting evaluations.

4. EVALUATION PRACTICES

Evaluations at USAID will be:

Integrated into design of projects: USAID's renewed focus on evaluation as a complementary and reinforcing relationship with other efforts to focus projects on achieving measurable results. These include a revival of project design capacity and strengthening the disciplinary expertise in priority areas, including stabilization, agriculture, economics and democratic governance. Compared to evaluations of projects with weak or vague causal maps and articulation of aims, we can expect to learn much more from evaluations of projects that are designed from the outset with clear development hypotheses, realistic expectations of the value and scale of results, and clear understanding of implementation risks.

For each project, consideration will be given during the design phase to the performance evaluation(s) and, in some cases, impact evaluation(s) that will be undertaken. This is part of the preparation of a Performance Management Plan. For Missions engaged in the preparation of a three- to five-year Country Development Cooperation Strategy, mission leadership will identify at least one opportunity for an impact evaluation for each Development Objective. Identifying key evaluation questions at the outset will both improve the quality of the project design, and will guide data collection during implementation.

When a project that will be subject to evaluation is initiated, baseline data, including variables that correspond to key outcomes and impacts, will be collected using high-quality methods and analyzed to establish a reference point. As a rule, baseline studies should collect sex-disaggregated data. To obtain baseline data, household or individual surveys are often valuable baseline data, and can be replicated toward the conclusion of implementation to assess changes.

Significant attention is required to ensure that baseline data are collected early in the project lifespan, before any significant implementation has occurred. In addition, the baseline data collection should be designed based on a plan for analysis of the data, to ensure that the appropriate variables are obtained and that, if probability sampling is used, the sample size is large enough to permit valid statistical comparisons.

Working closely with the responsible Program Office, project managers will ensure that implementing partners collect relevant monitoring data, and maintain data and documentation that can be accessed for future evaluations.

In cases where impact evaluations are undertaken to examine the relationship between an intervention or set of interventions and changes in a key development outcome, a parallel contractual or grant agreement will be established at the inception to accompany implementation. That contractual or grant agreement will include sufficient resources for data collection and analysis. Under unusual circumstances, when a separate arrangement is infeasible, implementing partners may subcontract an impact evaluation of a project subcomponent.

If and when opportunities exist to evaluate the impact of particular interventions, or to compare variations in implementation strategies, implementing partners are encouraged to bring these opportunities to the attention of the responsible technical officers. Technical officers can determine whether and how to support such impact evaluations to be conducted by the partner or, ideally, externally through a separate mechanism.

Unbiased in measurement and reporting: Evaluations of USAID projects will be undertaken so that they are not subject to the perception or reality of biased measurement or reporting due to conflict of interest or other factors. In most cases, the evaluations will be external (i.e., a third-party contractor or grantee, managed directly by USAID, will evaluate the projects), and the contract or grant for the evaluation will be managed by an operating unit's Program Office.

Whereas most evaluations will be external, funding may be dedicated within a project design for implementing partners to engage in evaluative work for their own institutional learning or accountability purposes. In cases where project funding from USAID supports an evaluation conducted or commissioned by an implementing partner, the findings from that evaluation must be shared in written form with the responsible technical officer within three months of the evaluation's conclusion.

In cases where USAID operating unit management determines that appropriate expertise exists within the Agency, and that engaging USAID staff in an evaluation will facilitate institutional learning, an evaluation team may be predominantly composed of USAID staff. However, an outside expert with appropriate skills and experience will be recruited to lead the team, mitigating the potential for conflict of interest. The outside expert may come from another U.S. Government Agency uninvolved in project implementation, or be engaged through a contractual mechanism.

Relevant: Evaluations will address the most important and relevant questions about project performance. In general, the importance and relevance will be achieved by explicitly linking evaluation questions to specific future decisions to be made by USAID leadership, partner governments and/or other key stakeholders. Those decisions frequently will be related to how resources should be allocated across and within sectors and thematic areas and/or how implementation should be modified to improve

effectiveness. To ensure relevance, consultation with in-country partners and beneficiaries is essential. Evaluation reports should include sufficient local and global contextual information so that the external validity and relevance of the evaluation can be assessed. Evaluations that are expected to influence resource allocation should include information on the cost structure and scalability of the intervention, as well as its effectiveness.

Based on the best methods: Evaluations will use methods that generate the highest quality and most credible evidence that corresponds to the questions being asked, taking into consideration time, budget and other practical considerations. Given the nature of development activities, both qualitative and quantitative methods yield valuable findings, and a combination of both often is optimal; observational, quasi-experimental and experimental designs all have their place. No single method will be privileged over others; rather, the selection of method or methods for a particular evaluation should principally consider the empirical strength of study design as well as the feasibility.

For impact evaluations, experimental methods generate the strongest evidence. Alternative methods should be utilized only when random assignment strategies are infeasible.

Evaluation methods should use sex-disaggregated data and incorporate attention to gender relations in all relevant areas. Methodological strengths and limitations will be communicated explicitly both in evaluation scopes of work and in evaluation reports.

Oriented toward reinforcing local capacity: The conduct of evaluations will be consistent with institutional aims of capacity building and respectful engagement with all partners. To the extent possible, evaluation specialists with appropriate expertise from partner countries, but not involved in project implementation, will lead and/or be included in evaluation teams. USAID will place priority within its sectoral programming on supporting partner government and civil society capacity to undertake evaluations and use the results generated.

Transparent: Findings from evaluations will be shared as widely as possible, with a commitment to full and active disclosure. Furthermore, a summary including a description of methods, key findings and recommendations will be available to the public on-line in a fully searchable form within three months of an evaluation's conclusion, as described below. Principled exceptions will be made in cases of classified or proprietary material.

5. EVALUATION REQUIREMENTS

Recognizing the diversity of project sizes and scopes across Agency activities, the application of evaluation requirements will occur at the level of the operating unit.

Evaluations of large projects and programs: Each operating unit is required to conduct at least one performance evaluation of each large project it implements. For these purposes, a "large project" is one that equals or exceeds in dollar value the mean (average) project size for the operating unit; this is intended to ensure that the majority of resources under management will be subject to evaluation. (In cases where there are factors that make it difficult to calculate mean project size—for example, when many projects are co-funded with other USG partners—operating units should consult with PPL/LER.)

The performance evaluation must be timed so that the findings will be available as decisions are made about new strategies, project designs and procurements. This will often mean, for example, that the evaluation will be designed and commissioned 18 months or more before the planned conclusion of the project.

Operating units are encouraged to identify opportunities for evaluations of their programs, rather than focusing only at the project level. Such evaluations are particularly valuable in the period preceding the

development of a new strategy, when questions are likely to be asked about the overall effectiveness of engagement in a particular sector or broad portfolio of activities.

Evaluations of innovative development interventions: In addition to the requirement above, any activity within a project involving untested hypotheses[2] or demonstrating new approaches that are anticipated to be expanded in scale or scope through US Government foreign assistance or other funding sources will, if feasible, undergo an impact evaluation. If it is not possible to effectively undertake an impact evaluation, operating units may instead undertake a performance evaluation, provided that the final evaluation report includes a concise but detailed statement about why an impact evaluation was not conducted. Regardless of whether an impact or performance evaluation is selected, the evaluation will be integrated into the design. Any activity or project designated as a "pilot" or "proof of concept" will fall under this requirement.

USAID operates in many environments where, due to security concerns, evaluations involving extensive site visits, interactions with beneficiaries, and other standard approaches are impossible. Moreover, even where security concerns are not paramount, some of the contexts in which USAID operates are so complex that standard linear and/or causal models may have little relevance. While this does not obviate the need for evaluations, creative and sometimes unorthodox approaches will be required to measure project achievements in complex and/or insecure environments. Using existing materials as a starting point, supplemental guidance and tools will be developed by PPL/LER, in collaboration with the relevant technical and regional bureaus, to support this work.

Evaluation Procedures: Evaluations will be undertaken in a manner that ensures credibility, unbiasedness, transparency, and the generation of high quality information and knowledge. Given the variation in evaluation questions and conditions, the means toward these ends will vary greatly from case to case. However, USAID evaluations of all types will use sound social science methods and include the following basic features:

• Establishment of team with the appropriate methodological and subject matter expertise to conduct an excellent evaluation.

• Written design, including identification of key question(s), methods, main features of data collection instruments, data analysis plans, and dissemination plan. Except in unusual circumstances, the design will be shared with country-level stakeholders as well as with the implementing partners for comment before being finalized.

• Gender-sensitive indicators and sex-disaggregated data.

• Approach that encourages participation by national counterparts and evaluators in the design and conduct of evaluations.

• Use of data collection and analytic methods that ensure, to the maximum extent possible, that if a different, well-qualified evaluator were to undertake the same evaluation, he or she would arrive at the same or similar findings and conclusions.

• Application and use to the maximum extent possible of social science methods and tools that reduce the need for evaluator-specific judgments.

• Standardized recording and maintenance of records from the evaluation (e.g., focus group transcripts).

---

[2] Whether a hypothesis is "tested" or "untested" is often a matter of professional judgment. However, in the project design phase an effort should be made to synthesize the best available evidence regarding the intervention(s) being included in the project – for example, the approach to teacher training, the use of performance incentives to improve health worker productivity, or the strategy to foster community development through strengthening local governance bodies. Where a truly novel approach is being introduced and there is little or no empirical evidence regarding its effectiveness in any setting, this would be characterized as "untested."

- Collection of data on variables corresponding to inputs, outputs, outcomes and impacts, as well as financial data that permits computation of unit costs and analysis of cost structure.
- Evaluation findings that are based on facts, evidence and data. This precludes relying exclusively upon anecdotes, hearsay and unverified opinions. Findings should be specific, concise and supported by quantitative and qualitative information that is reliable, valid and generalizable.
- Evaluation reports that include the original scope of work, a full description of methodology (or methodologies) used, as well as the limitations in the inferences that can be drawn. Readers should have sufficient information about the body of evidence and how information was gathered to make a judgment as to its reliability, validity and generalizability.
- Evaluation reports that include action-oriented, practical and specific recommendations assigning or designating the implementer.
- Evaluation reports that are shared widely and in an accessible form with all partners and stakeholders, and with the general public.
- Adequate budget and timeline for a high quality evaluation.

    To assure evaluation quality, the following systems will be put into place:

- Scopes of work for evaluations shall include criteria for the quality of the evaluation report. These are provided in Appendix 1.
- The operating unit Program Office will organize peer reviews of evaluation scopes of work and draft evaluation reports, seeking support from the corresponding regional and technical bureaus.
- PPL/LER will organize occasional external technical audits of operating units' compliance with the evaluation policy, including through random checks of scopes of work, evaluation reports and utilization of evaluation findings.

    Evaluation Transparency: The presumption of openness in the conduct of USAID evaluations will be manifested at two stages: (1) when an evaluation design is agreed upon; and (2) when the evaluation report has been completed, that report will be disseminated. Compliance will include:

    Registration: Operating units will provide information on-line in a fully searchable form about the initiation of evaluations and expected timing of release of findings. This information will be included in the annual PPR, and communicated to the public on the USAID website.

    Standard Reporting and Dissemination: In addition to the findings and methodology documented in each evaluation report, other key characteristics of each report include:

- Disclosure of conflict of interest: For external evaluations, all evaluation team members will provide a signed statement attesting to a lack of conflict of interest, or describing an existing conflict of interest relative to the project being evaluated.
- Statement of differences: When applicable, evaluation reports will include statements regarding any significant unresolved differences of opinion on the part of funders, implementers and/or members of the evaluation team.

    Completed evaluations (and drafts of reports of evaluations completed more than three months prior) must be submitted to the DEC and a cover sheet attached indicating the type of evaluation conducted and design. Each completed evaluation must include a 3- to 5-page summary of the purpose, background of the project, main evaluation questions, methods, findings, conclusions, recommendations and lessons learned (if applicable) of the evaluation.

    Data Warehousing: All quantitative data collected by USAID or one of the Agency's contractors or grantees for the purposes of an evaluation must be uploaded and stored in a central database.

(Guidance regarding the procedures for this warehousing will be issued at a later date.) The data should be organized and fully documented for use by those not fully familiar with the project or the evaluation.

Utilization of Evaluation Findings: Evaluation is useful only insofar as it provides evidence to inform real-world decision making. Every step of USAID's programming model – from design to implementation to evaluation – will be undertaken from the perspective not only of achieving development objectives, but of contributing to the broader goal of learning from experience. The learning from previous experience that is captured in evaluation findings should be easy to access and considered whenever an officer is designing and implementing new projects, and project and policies should be designed so they are evaluable (when possible) and should include a plan for evaluation. The utilization of evaluation findings will be encouraged in the guidance in Mission Orders, and will be highlighted in Country Development Cooperation Strategies. In addition, PPL/LER will commission occasional external technical audits to determine whether and how evaluation findings are being used for decision making by operating units.

Evaluation Resources: USAID recognizes that evaluation findings have significant value to the institution's effectiveness, and merit adequate resources. Additional resources at the central level, for training, technical support, quality control, and guideline development help to leverage the investments currently being made in evaluation throughout the Agency. Moreover, new resources – primarily in the form of qualified professional staff at the mission and regional missions – and access to technical support through indefinite quantity contracts and other flexible mechanisms are also needed.

Additional Human Resource Development and Staff: Explicit competencies for evaluation specialists and senior evaluation specialists will be developed by PPL/LER, integrated into human resource policies and practices, and updated as needed. These competencies will reflect the skill sets required to implement this policy. One or more certificate courses will be offered to enhance the skill set of existing staff. In addition, the Office of Human Resources, with input from PPL/LER will determine the complement of evaluation specialists required within the staffing model to fulfill the needs of policy implementation. It is anticipated that this will require hiring and/or redeployment of evaluation specialists and senior evaluation specialists.

Procurement Mechanisms for Evaluation Services: Implementation of this policy will increase the demand for highly trained and experienced evaluation specialists.   In particular, indefinite quantity contracts focusing on particular thematic areas and/or methodologies may, when appropriate, be used as mechanisms to ensure timely access to specialist services of high quality. Country and regional missions, as well as technical bureaus, are encouraged to develop procurement mechanisms that permit timely access to appropriate evaluation expertise.

Financial Resources: USAID will devote approximately 3 percent of total program dollars, on average, to external performance and impact evaluation. This is distinct from resources dedicated to monitoring. In some instances, this may require reallocation away from project implementation, particularly when the opportunity to improve effectiveness through learning is deemed to be very large. In addition, USAID acknowledges that more intensive evaluation efforts may increase the need for dedicated monitoring and data collection resources within contracts and grants to implementing partners.

6. CONCLUSION

USAID's ability to fulfill commitments to accountability and obtain the benefits of institutional learning depends, in large part, on embedding excellent evaluation practices throughout the organization. No single policy can anticipate and provide detailed guidance for the diverse set of USAID projects and contexts. However, this policy seeks to establish the roles and responsibilities, and the key expectations regarding the design, conduct, dissemination and use of evaluation. While the policy has an indefinite

term, we expect that as it is implemented, new and better ideas will emerge about how to improve evaluation and make it more relevant to USAID's institutional environment. Over time, those ideas will be integrated into the Agency's work through an update of this policy.

Appendix 1

Criteria to Ensure the Quality of the Evaluation Report

- The evaluation report should represent a thoughtful, well-researched and well organized effort to objectively evaluate what worked in the project, what did not and why.
- Evaluation reports shall address all evaluation questions included in the scope of work.
- The evaluation report should include the scope of work as an annex. All modifications to the scope of work, whether in technical requirements, evaluation questions, evaluation team composition, methodology or timeline need to be agreed upon in writing by the technical officer.
- Evaluation methodology shall be explained in detail and all tools used in conducting the evaluation such as questionnaires, checklists and discussion guides will be included in an Annex in the final report.
- Evaluation findings will assess outcomes and impact on males and females.
- Limitations to the evaluation shall be disclosed in the report, with particular attention to the limitations associated with the evaluation methodology (selection bias, recall bias, unobservable differences between comparator groups, etc.).
- Evaluation findings should be presented as analyzed facts, evidence and data and not based on anecdotes, hearsay or the compilation of people's opinions. Findings should be specific, concise and supported by strong quantitative or qualitative evidence.
- Sources of information need to be properly identified and listed in an annex.
- Recommendations need to be supported by a specific set of findings.
- Recommendations should be action-oriented, practical and specific, with defined responsibility for the action.