

U.S. DEPARTMENT OF STATE

Department of State Program Evaluation Policy

February 2012

A. INTRODUCTION

The Department of State is the lead institution for the conduct of American diplomacy and all foreign affairs efforts are paid for by the foreign affairs budget. The Department is committed to using performance management best practices, including where feasible and useful, program evaluation, to achieve the most effective U.S. foreign policy outcomes and greater accountability to our primary stakeholders, the American people. This Department policy embodies the aims of the current National Security Strategy, the Presidential Policy Directive on Global Development, and the Quadrennial Diplomacy and Development Review. The policy applies to evaluations of the Department's diplomatic and development programs, projects and activities.

A robust, coordinated and targeted evaluation function is essential to the ability of the Department to measure and monitor program performance; make decisions for programmatic adjustments and changes; document program impact; identify best practices and lessons learned; help assess return on investment; provide inputs for policy, planning and budget decisions; and assure accountability to the American people.

With the passage of the Government Performance and Results Act (GPRA) of 1993 and the GPRA Modernization Act (GPRAMA) of 2010, Congress strengthened the mandate to evaluate programs and required agencies to include a discussion of program evaluations in their strategic plans and performance reports. Although USAID and the Department have made significant progress toward fulfilling these mandates, this policy is intended to provide a more coherent evaluation policy and a coordinated strategy and work plan for conducting evaluations in the Department consistent with GPRAMA, the Clinger-Cohen Act of 1996, the E-Government Act of 2002 and OMB Memorandum M-11-29, Chief Information Officer Authorities.

This policy also responds to today's needs. Many stakeholders are demanding greater transparency in decision-making and disclosure of information.

This policy works in concert with existing and pending Department policies, strategies and operational guidance, including those that are forthcoming from the QDDR implementation effort. In addition, the policy is consistent with the USAID evaluation policy. The Department will continue to work collaboratively with USAID and other USG agencies to ensure that the organizations' guidelines and procedures with respect to evaluation are mutually reinforcing.

Key Definitions

Evaluation: 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.

Performance Evaluation: Performance evaluations focus on the implementation of an intervention (or management effort) and describe 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.

Impact Evaluation: Impact evaluations assess/measure the change in the targeted groups, organizations and other beneficiaries that can be attributed to a project, program or policy. 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. The essence of impact evaluation lies in establishing that the changes have occurred as a result of the intervention, or at least the latter has substantially contributed to it. Impact evaluations require defining a counterfactual by establishing a treatment and a control group.

Summative/ Ex-Post Evaluations: They differ from performance evaluations in that their focus is primarily on outcomes and impacts but they often include evaluation of effectiveness and are conducted when an intervention has ended or is soon likely to end.

Global/Regional Evaluations: Are designed to examine the performance and outcomes of a major sector or sub-sector of foreign affairs programs to draw general findings, conclusions and lessons. The purpose of a global evaluation is not to evaluate the success or failure of individual projects but to determine the efficacy and outcomes of the programs per se.

Experience Reviews: These involve systematic analysis of past experience. They focus on a limited range of questions that are of paramount interest to policy and decision makers. Experience reviews do not require fieldwork, and the needed information is gathered from sources such as literature reviews, experts' workshops, key informant interviews and/or surveys.

Monitoring: Monitoring provides an indication of progress against goals and indicators of performance, reveals whether desired results are occurring, and confirms whether implementation is on track. In general the results measured are the direct and near term consequences of program activities; whereas evaluations document the achievement of outcomes and results and, in some cases, the value of continuing the investment.

Performance Indicator: Performance indicators measure a particular characteristic or dimension of an intervention's outputs or outcomes. Outputs are directly attributable to the intervention activities, while outcomes represent results to which a given program contributes but for which it is not solely responsible.

Program: A program is a set of activities or projects that are typically implemented by several parties over a specified period of time. A program evaluation often involves an assessment or study of multiple activities that may cut across sectors and/or geographic areas.

Project: A project is a set of planned and then executed interventions identified through a design process, which are together intended to achieve a defined 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. A set of projects makes up the portfolio of a program. A project evaluation is often carried out within the framework of a broader program.

Activity: An activity is a component of a project that contributes to a project purpose. It refers to an award (such as a contract, grant or cooperative agreement), or a component of a project such as training or technical assistance.

B. PURPOSE

This new evaluation policy proposes a framework, where applicable, to implement evaluations of programs, projects, and activities that are carried out and funded by the Department. This policy is intended to provide clarity about the purposes of evaluation, the requirements for evaluation, the types of evaluations that are possible, and an approach for conducting, disseminating and using evaluations.

Evaluation at the Department has two primary purposes: (1) accountability to stakeholders and (2) learning, which could inform and improve program, project, and activity design and implementation.

Accountability: Well designed and timely evaluations help to ensure accountability for the USG resources spent on foreign affairs activities. Evaluations enable program managers and leadership to determine the cost effectiveness of programs, projects, initiatives, activities, interventions, etc., and in the case of a program or project, quality of its planning and implementation. Consequently, evaluation findings can provide empirical data for reports to various stakeholders in foreign assistance planning and in larger diplomatic activities.

Learning: Evaluations document the results, impact, or effectiveness of organizational and programming activities, thereby facilitating learning from experience. The Department can apply such learning to the development of new projects, programs, strategies and policies. Empirically-grounded evaluations also aid informed decision making when considering new programs or projects, interventions, activities, etc.

C. APPLICABILITY

Effective February 15, 2012 this policy applies to all State Bureaus and S/GAC. Other offices designated by the Performance Improvement Officer (PIO) may also be requested to adhere to this policy on a case by case basis.

D. EVALUATION REQUIREMENT

The Department of State requires that all large programs, projects, and activities be evaluated at least once in their lifetime or every five years, whichever is less. Bureaus will determine which programs, projects or activities to evaluate; upon request RM and F will assist in this effort. For most bureaus, a "large" program, project, or activity is one whose dollar value equals or exceeds the median program/ project/ activity size for the bureau. For bureaus that conduct their programs, projects, and activities primarily with staff resources, a "large" program, project, or activity can instead be identified using criteria based on the number of full time staff (or equivalent). In this case, a "large" program, project, or activity is one for which the number of full time staff associated with the program, project, or activity exceeds the

median number of staff associated with individual programs, projects, and activities in the bureau. Noting the diversity of undertakings at State, this policy refers broadly to programs, projects, and activities—such as diplomatic functions-- so that bureaus will identify the level at which most of their funding is used and apply the policy accordingly, making sure that significant undertakings are evaluated regardless of whether they are programs, projects, or activities. Similarly, program “lifetime” will be defined by bureaus based on the implementation period of various programs within their portfolios. These criteria are intended to ensure that the most important initiatives under stewardship by the bureau are subject to evaluation. Keeping these guidelines in mind, bureaus have the flexibility to determine what constitutes a large program or activity in the context of their portfolio. Bureaus will be responsible for determining the selection, timing, process and manner of evaluation for specific programs and projects. Additionally, RM and F, in consultation with bureaus, may identify other programs/projects for evaluation.

The Department also requires that all pilot programs, whether funded with Foreign Assistance or State Operations monies, include a description of the parameters to be studied, how such parameters will be measured, and by whom they will be measured; as well as the planned disposition of the results after the pilot program ends. Pilot programs should be designed to run only long enough to provide meaningful results, and must not be designed to circumvent rulemaking. Pilot programs and projects designed and executed to directly achieve a critical foreign policy objective should be evaluated once every five years.

Each Bureau is required to evaluate two to four projects/programs/activities over the 24 month period beginning with FY 2012, depending on the size, scope and complexity of the programs/projects being evaluated and funding availability. In FY 2013 this requirement will be extended to posts. The Department and USAID will coordinate on the implementation of both agencies’ evaluation policies at the mission and post level and will do so in concert with existing strategic planning processes.

It is important to note that this policy does not alter existing monitoring and performance procedures for grants and contracts, as suggested by the Office of the Director of U.S. Foreign Assistance (F) and the Bureau of Administration (A).

Minimizing Duplication of Effort: This policy recognizes that Bureaus do not always directly implement a project, program or activity and clarifies the policy in those cases. In many cases Bureaus provide funds to other agencies, operating units, or, particularly in the case of foreign assistance, other international organizations. In such cases Bureaus are expected to ensure that the implementing organization carries out evaluations of programs, projects, and activities consistent with these guidelines. Bureaus may count evaluations carried out by implementing partners toward the requirement to evaluate two to four projects/programs/activities within 24 months. In addition Bureaus should work with organizations to assure that duplicative evaluations are not required of programs, projects, or other joint activities.

E. TYPES OF EVALUATIONS

When choosing the type of evaluation for a program, project, or activity, the bureau will need to decide whether to conduct a performance evaluation or an impact evaluation. Performance evaluations examine the inputs, outputs, outcomes, and performance of an intervention; whereas impact evaluations focus on measuring the change attributable to a given intervention. Bureaus must decide the most appropriate type of evaluation based on the specific program, project, or activity but it is expected that performance evaluations will comprise the great majority of the

evaluations performed because impact evaluations are more time consuming, costly, and often difficult to successfully design for State programs, projects, and activities. Regardless of the type of evaluation, all evaluations must be context-sensitive and methodologically sound.

GAO and OIG studies and audits do not qualify as evaluations for the purposes of this policy but bureaus are encouraged to use such studies to inform the planning of evaluations, as applicable.

F. THE ROLE OF THE OFFICE OF U.S. FOREIGN ASSISTANCE RESOURCES (F) AND THE BUREAU OF RESOURCE MANAGEMENT (RM)

The primary role of F and RM is to monitor and report annually to the Deputy Secretary of State for Management progress each bureau has made on their Bureau Evaluation Plan. F and RM will also feature key findings in an annual report to the Secretary. While it is not the primary role of F and RM to determine which programs, projects, activities, etc., that will be evaluated, RM and F may recommend programs, projects, etc., for evaluation and work with bureaus to determine appropriate evaluative approaches. Finally, RM and F will offer support in implementing the requirements of this policy. See Section H for further details.

G. EVALUATION STANDARDS

Methodological Rigor: Evaluations should be “evidence based,” meaning that they should be based on verifiable data and information that have been gathered using the standards of professional evaluation organizations. Such data can be both qualitative and quantitative. Bureaus can use a wide variety of evaluation designs such as case studies, experimental and quasi experimental designs as well as data collection methods such as key informant interviews, surveys, direct observation or focus group discussions. Regardless, such data must pass the tests of reliability and validity which are different for qualitative and quantitative data.

Independence and Integrity: To the extent possible, bureaus should ensure that the evaluators are free from any pressure and/or bureaucratic interference. Managers of the programs, projects or activities being evaluated should not be in a position to influence evaluation findings and recommendations. When evaluators are supervised by the manager of the program being evaluated, they may be susceptible to his/her influence and reluctant to highlight intervention shortcomings that may reflect poorly upon the program, project, or activity being evaluated. Such a structure may also create the appearance of a conflict of interest, creating doubts about the objectivity of the evaluator and the credibility of the evaluation. Independence does not, however, imply isolation from project/program managers. In fact, active engagement of bureau staff and managers is necessary to conduct an evaluation as long as they are not in a position to improperly influence the outcome. This same requirement also applies to contractors and grantees.

H. EVALUATION RESPONSIBILITY

Agency Level: The Bureau of Resource Management (RM) and the Office of the Director of Foreign Assistance Resources (F) will work closely with the Performance Improvement Officer to assist bureaus in implementing the evaluation policy. RM will serve as the coordinator for evaluations of State Operations-funded programs, projects, and activities. F will serve as the coordinator for evaluations of Foreign Assistance-funded programs, projects, and activities.

RM and F will work together to coordinate tools, technical support, funding, and evaluation training to assist bureaus in the implementation of this policy. Support activities will include staff training, the implementation of Evaluation Indefinite Delivery, Indefinite Quantity (IDIQ) contracts, the establishment of a State evaluation Community of Practice, and the publication and dissemination of a toolkit of evaluation standards, guidelines, and a glossary of terms. Technical assistance will also be available to assist in the development of appropriate scopes of work, evaluation designs, evaluation methodologies, analysis and reporting.

Together, RM and F will prepare an annual report for the Secretary that highlights key evaluation practices and findings, as well as changes and challenges in evaluation practice at State. Information for this report will come from completed evaluation reports, Bureau Evaluation Plans and the Performance Plan and Report (PPR).

Bureau Level: The evaluation function shall have two levels of responsibility:

- Management: The Bureau allocating funds for a specific program, project or activity is responsible for ensuring that an evaluation is planned, budgeted and conducted.
- Implementation: The Bureau implementing and overseeing a program, project, or activity is responsible for implementing and managing its evaluation.

As noted above, training will address bureau level responsibilities and Department level activities in support of the policy.

I. EVALUATION USE

Evaluations are only meaningful if they are used. Properly designed evaluations will provide insights and evidence to inform real-world decision making and contribute to learning. Bureaus therefore should integrate evaluation findings into decision-making about strategies, program priorities and project design, as well as into planning and budget formulation processes. For example, evaluation findings should be used to course-correct in interim years of a bureau's multi-year strategic plan as necessary, or to shape that plan initially. RM and F will provide guidance to assist bureaus to integrate evaluation findings into decision making about strategies, program priorities, and project design and resource requests.

J. BUREAU EVALUATION PLANS

All bureaus are required to develop and submit a Bureau Evaluation Plan (BEP) as an annex to their multi-year strategic plan. The BEP must, at a minimum cover a three year period. After initial submission Bureaus will need to update their BEPs annually. Bureaus responsible for foreign assistance activities will submit annual updates to the BEP in the PPR (including information on any evaluations for state operations activities). Bureaus solely responsible for State Operations activities will report annual updates to the BEP in a parallel data call. Bureau Evaluation Plans need to encompass new programs and major policy initiatives—such as cross-cutting, collaborative efforts—as well as existing programs and projects. The BEP should also include evaluations that are planned by implementing organizations in cases where Bureaus provide funds to other agencies, operating units, or international organizations as outlined in section D above.

Bureau Evaluation Plans will be submitted annually to the PIO through RM and F. Plans will be reviewed by this group and feedback provided to each Bureau.

K. BUREAU POINT OF CONTACT

Each bureau must identify a senior staff person to serve as the bureau Point of Contact to ensure that the evaluation function is fully operational and integrated into the planning and decision making process. The coordinator shall be a Deputy Assistant Secretary (DAS) or his/her designee. He or she will serve as the main point of contact in the bureau on evaluation and will interact with the PIO, RM and F on the bureau's evaluation efforts and compliance with this policy. The time and staff allotted to this function should be commensurate with the size of the evaluation portfolio being managed and should be included in the bureau budget.

L. EVALUATION RESOURCES

Performance monitoring and evaluation are allowable program costs. The cost of an evaluation will vary widely by program, and no set amount is prescribed. International professional standards generally set a range between 3% and 5% of total program, labor or project costs. Fully recognizing that the Department and bureaus face funding constraints, program managers should identify resources of up to 3-5% for evaluation activities.

M. DOCUMENTING AND SHARING EVALUATIONS REPORTS

Bureaus and posts shall maintain an official copy of completed evaluation reports and must also electronically transmit final evaluation reports as a cable using the "RVAL" TAG so that they can be collected into a central repository. Evaluations that are unclassified, but warrant/require administrative control and protection from public or other unauthorized disclosure, should be marked SBU in accordance with 12 FAM 541. Classified evaluations must be marked with appropriate classification markings.

In keeping with the policy of USG transparency, bureaus and posts are required to post copies of their evaluation reports on their OpenNet or ClassNet website homepage so that other State Department colleagues can benefit from them and so that evaluation results can facilitate learning within the Department. In addition, schedules for future evaluations will be included in the State-USAID Joint Strategic Plan, and summaries of completed evaluations will be reported in the Department's annual performance report.