

Guidance for the Design, Monitoring and Evaluation Policy at the Department of State



January 2019 Edition

Contents

1. PURPOSE	4
2. TIMELINE FOR COMPLIANCE WITH 18 FAM 300	4
3. IDENTIFYING MAJOR PROGRAMS AND PROJECTS.....	4
4. PROGRAM/PROJECT DESIGN	5
Programs.....	7
Projects	7
5. MONITORING	7
Baseline data.....	8
Performance Management Plans	9
6. EVALUATION	9
Evaluation Requirements.....	11
Bureau Evaluation Coordinators.....	12
Bureau Evaluation Plans	13
Considerations for Evaluation	13
Ethical Standards	14
Types of Evaluation.....	15
Collaboration and Other Ways to Evaluate	17
Planning and Managing Evaluations.....	20
Specifying Objectives and Audience.....	20
Formulating Evaluation Questions	20
Determining Data Collection Methods.....	21
Selecting an Evaluation Design.....	23
Preparing a Statement of Work.....	24
Contracting Evaluations.....	25
Evaluation Report Review and Follow up.....	27
Technical Assistance and Capacity Building	28
Evaluation Use	28
Dissemination	29

Foreign Assistance	29
Diplomatic Engagement	30
7. Analysis and Learning	30
8. Implementation	30
Budgeting for Program Design, Monitoring, and Evaluation Activities.....	31
Transfers of Funds	32



1. PURPOSE

This guidance provides direction to managers in bureaus and independent offices at the Department of State on how to execute the requirements of 18 FAM 300, which involves identifying, designing, monitoring, evaluating, and assessing the progress of Department programs and projects. Consistent with 18 FAM 300, the purpose of this guidance is to clarify the ways in which Department bureaus and independent offices can establish a clear line of sight from *what* they want to achieve as documented in strategic plans, to *how* they intend to achieve it through key programs, projects, or processes, and how they will determine if those efforts are *working as intended* based on monitoring, evaluation, and learning activities. The steps in this guidance can be applied to large programs with no defined timeframe, as well as to projects with defined timeframes, resources, and deliverables.

2. TIMELINE FOR COMPLIANCE WITH 18 FAM 300

The Office of U.S. Foreign Assistance Resources (F) and the Bureau for Budget and Planning (BP) can offer technical assistance to bureaus and independent offices to execute 18 FAM 300.

June 29, 2018: In consultation with F or BP, bureaus and independent offices have *identified* their major programs and projects per section 3 of this guidance.

March 29, 2019: For all of their major programs and projects, bureaus and independent offices have completed all of the design steps identified in section 4 of this guidance.

June 28, 2019: For all of their major programs and projects, bureaus and independent offices have established monitoring and evaluation plans that identify relevant indicators, and any possible opportunities for evaluation, if applicable, per sections 5 and 6 of this guidance.

Ongoing: Bureaus and independent offices enter all planned, ongoing, and completed evaluations into the Evaluation Registry or EMS, per section 6 of this guidance.

Ongoing: Bureaus and independent offices assess progress and results, and use that information to inform management decisions, per section 7 of this guidance.

Ongoing: OIG uses this timeline to assess bureaus' compliance with 18 FAM 300.

3. IDENTIFYING MAJOR PROGRAMS AND PROJECTS

All bureaus and independent offices must identify the major programs and projects they undertake to achieve their broader goals, as this sets the foundation for the remaining design, monitoring, and evaluation requirements in 18 FAM 300. The identification process should include reviewing the relevant strategic plan(s), authorities, Executive Orders, funding streams, cross cutting issues or themes the bureau or independent office is responsible for,

organizational structure, or countries in which the bureau or independent office has equities. Ultimately, the bureau or independent office must define its major programs and projects in a way that makes it possible to develop a logic model or project charter for them specifying their key inputs, activities, outputs, and expected outcomes. Monitoring plans, evaluation plans, and learning activities will then track back to assessing the extent to which the overall logic and “theory of change” the bureau or independent office has articulated for the program or project is progressing as planned.

4. PROGRAM/PROJECT DESIGN

Design is the process of analyzing the context, identifying the root causes of issues to be addressed, and constructing logic and a theory of how and why a program or project is expected to work. Design activities may be conducted by bureau or independent office staff, grantees, contractors, or other U.S. government agencies to which Department of State funds are transferred. All documents and information should be maintained in bureau or independent office files.

The Department of State Program Design and Performance Management (PD/PM) Toolkit provides procedures and templates for designing programs, projects, and processes, as detailed below. Bureaus and independent offices may use alternative approaches or templates to execute the required design steps. Minimum standards that must be met for design are detailed in the Program and Project sub-sections following Table 1.

We recommend that language be added to any grant or award to an implementer to clarify that if requested, the implementer will cooperate with and facilitate the work of an independent contractor commissioned to design and implement an evaluation of the program.

Table 1: Design Requirements and Corresponding PD/PM Toolkit Sections

Required Design Step	Purpose	Applies To	PD/PM Toolkit Section
Alignment to Strategy	Assess how a program/project idea can best align with and advance existing strategies or other higher level directives.	Programs Projects	1
Internal Assessment	Understand the capabilities within your own bureau, office, or post.	Programs Projects	2
External Assessment	Survey the context where the program will take place to establish baseline information and understand opportunities and threats.	Programs Projects	

	Supporting evidence and information from third parties may be used, such as research institutions, country government, civil society groups, etc.		
Root Cause Analysis	Analyze the source of the problem or issue giving rise to the program/project.	Programs Projects	
Review of the Issue	Identify the problem to be solved or issue to be addressed and connecting it to its stakeholder, context and beneficiaries.	Programs Projects	
Problem/Need Statement	Create a clear description of the problem or issue to be addressed and establish program/project focus.	Programs Projects	3
Goals and Objectives	Define the broad aims of the program or project and more specific achievements to be realized over the course of the program.	Programs Projects	
Logic Model and Theory of Change	A visual representation of the relationship between program inputs, activities, and short- and long-term outcomes. It should be accompanied by a theory of change summarizing why, based on available evidence, the changes described in the logic model are expected to occur.	Programs Projects*	
Program Assumptions	Identify the conditions that need to exist in order for one step of the logic model to succeed and lead to the next step.	Programs	
Scope	Outlines the project's deliverables and identifies the constraints, assumptions and key success factors. A well-written scope statement clearly defines the boundaries of a project	Project	Appendix A
Stakeholders	An individual, group, or organization, who may affect, be affected by, or perceive itself to be affected by a decision, activity, or outcome of a project	Project	Appendix A
Key Deliverables	Products or services provided to the project recipient that satisfies a milestone or due date	Project	Appendix A
Milestone Schedule	Dates that mark specific points along a project timeline	Project	Appendix A

*Projects may create a logic model or project charter, depending on the nature of the project.

Programs

Templates from Sections 1 through 3 of the Program Design and Performance Management Toolkit as identified in Table 1 should be completed to execute the design steps for bureau or independent office major programs, or bureaus or independent offices can complete comparable documents correlating to each required design step outlined in Table 1. Note that all logic models must, at a minimum, identify the program's inputs, activities, outputs, and outcomes and have a corresponding theory of change.

Projects

To document the design of major projects and satisfy the requirements of 18 FAM 300, bureaus and independent offices may use the templates provided in the Program Design and Performance Management Toolkit noted in Table 1, or they may complete a project charter. At a minimum, a project charter should contain:

- Goals. Identify one or more project goals that articulate the type of change that should occur as a result of the effort.
- Project scope. Determine what the project will cover and what it will not cover.
- Stakeholders. Identify the personnel who will be managing and supporting the project as well as those who will be providing input or will benefit from its outputs.
- Key deliverables. List key tasks and deliverables for the project
- Schedule. Identify due dates for milestones within the project as well as a project completion date.

The [Teamwork@State](#) website¹ has several relevant templates and instructions for project design, including a project charter.

5. MONITORING

Monitoring involves ongoing data collection against key performance indicators or milestones to gauge the direct and near-term effects of program activities and whether desired results are occurring as expected during program implementation. Monitoring data describe *what* is happening and inform whether implementation is on track or if any timely corrections or adjustments may be needed to improve efficiency or effectiveness. Monitoring data can also

¹ This link is to an internal website. Most links in this document, which is being published for the first time, are internal and will not be accessible until the guidance is next updated to conform to the 2020 update of the Department's Design, Monitoring, and Evaluation Policy, 18 FAM 300.

indicate when an evaluation is needed to understand how or why certain results are being observed, and can provide useful inputs into an evaluation.

Bureaus should refer to the Program Design and Performance Management Toolkit for guidelines on different types of indicators, and how to develop indicators that will provide useful information. The Toolkit also provides helpful templates and case examples to assist in developing a monitoring plan.

Table 2: Monitoring Requirements and Corresponding PD/PM Toolkit Sections

Monitoring Step	Purpose	Applies To	Toolkit Section
Indicators or Milestones	Indicators or milestones should be developed that directly correspond to the logic model. Often, indicators help track the extent to which activities, outputs, and some near-term outcomes are occurring. (Evaluations are often needed to assess the extent to which short- and long-term outcomes are occurring.)	Programs Projects	4.2
Indicator Reference Sheets	A description of each indicator the bureau or independent office is tracking to ensure it is used consistently by all stakeholders, and work through data collection feasibility considerations. Should be updated as necessary.	Programs Projects	4.2

All programs and projects should be monitored as described in the [Program Design and Performance Management Toolkit](#), Section 4.

Baseline data

Bureaus and independent offices are required to document the data collection methodology for their major programs and projects. While baseline data collection is ideally conducted before a program or project begins so that observed changes during or after implementation can be compared against the baseline, this is not always feasible or necessary. In particular, fluid environments and suitably adaptive activities (such as diplomacy) can make baselines measured at inception unnecessary, or amenable to alternative approaches. These include baseline collection that starts after a program or project begins, but far enough from the end that changes can still be measured; retroactive baseline estimates; and established methods that are alternatives to rigid pre- and post-program data comparisons, such as most significant

change² or outcome harvesting.³ Bureaus and independent offices must consider what is feasible for the particular circumstances of the program or project, and document the methodology to be used, and why.

For programs or projects that take place in multiple countries or regions, it is not necessary to use the same baseline for every country or region.

Performance Management Plans

Bureaus and independent offices should maintain performance management plans for all programs and projects as defined in Sections 4 and 5 of the [Program Design and Performance Management Toolkit](#).

6. EVALUATION

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 for federal agencies to evaluate their activities and programs and required them to include a discussion of evaluations in their strategic plans and performance reports. In addition, the Foreign Assistance Transparency and Accountability Act of 2016 (FATAA) set evaluation within performance management, emphasizing planning and design as well as transparency through public availability of evaluation results and other program information. This section provides further clarity about the purposes and requirements for evaluation, the types of evaluations the Department recommends, and the approach for conducting, disseminating, and using evaluations.

Monitoring and evaluation are often mentioned together as signified by the acronym M&E. While the two complement each other, they are conceptually and operationally different. *Monitoring* is a continual process designed to give an indication of progress against goals and indicators of performance, 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. Ideally, monitoring should be built into every project, activity, and program. *Evaluation*, on the other hand, is much more comprehensive. While it will make use of monitoring data and other evidence, evaluation goes beyond tracking progress to identifying the underlying factors and forces that affect the implementation process, efficiency, sustainability, and effectiveness of programs, projects, or processes.

² See, for example, Rick Davies and Jess Dart, “[The ‘Most Significant Change’ \(MSC\) Technique: A Guide to Its Use.](#)”

³ See, for example, Ricardo Wilson-Grau and Heather Britt, “[Outcome Harvesting.](#)”

There is an important difference between audits (such as those conducted by the OIG) and evaluations. The intent of an evaluation is for the commissioning bureau to *learn* how it can improve its program, project, activity, or process. The bureau is instrumental in formulating questions it wants answered and in managing the evaluation. On the other hand, an audit is typically externally commissioned with a focus on compliance, waste, fraud, or abuse. A bureau subject to an audit has no input to the questions asked or the focus of the work. Audits typically focus on whether or not programs and processes comply with existing or pre-defined standards and legal requirements whereas evaluations typically focus on questions a bureau wants answered about whether a program or process is appropriate, effective, or sustainable given existing conditions.

With this policy, the Department encourages bureaus and offices to ask questions about their programs, projects, and processes and answer them through evaluation to make improvements and facilitate better decision-making. As a result, audits commissioned and executed by external parties for purposes of their own cannot be counted toward the evaluation requirement.

Well-designed and empirically grounded evaluations should promote the following distinct but interrelated objectives:

1) **Learning:** Evaluations document the results, impact, or effectiveness of organizational and programming activities, thereby facilitating learning from experience. They produce knowledge – insights, information and recommendations – that can be used to improve the management of services as well as the implementation of existing programs. Such knowledge can also facilitate the development of new programs that are more efficient and effective.

2) **Improved Performance:** Most evaluations are undertaken to improve the performance of a program, project, or process by generating knowledge, information, and recommendations that can be used by their managers. For example, a mid-term evaluation of a civil society program in an African country can tell program managers why it is not reaching the targeted civil society organizations, what concerns and reservations these organizations have about the program, and finally what can be done to solve the problem.

3) **Informed Decision-Making:** The Department's officials encounter problems that require objective, timely information for decision making. They have to decide whether to continue or close an existing facility, drop programs that are not producing the expected results, adjust the scope of an existing activity, or develop new initiatives to address an unexpected challenge. For example, a bureau has to decide whether it should continue to invest heavily in civil society or has to make a decision about the future allocation of resources to a developing country. In both

cases, evaluations can aid informed decision-making. By commissioning an evaluation, concerned officials will be able to gather solid data and decide on a future course of action on the basis of the data and independent analysis rather than of hunches or professional opinion.

4) *Transparency and Accountability:* Well-designed, timely evaluations help to ensure accountability for USG resources spent on foreign affairs. Evaluations enable managers to determine the cost effectiveness of the Department's efforts and programming, as well as the quality of its planning and implementation. Consequently, evaluation findings provide empirical data for reports submitted to various stakeholders – Department managers, the Office of Management and Budget, the White House, the Congress, the NGO community, beneficiaries of State-funded programs, and above all, the American taxpayers.

5) *Engagement, Ownership, and Capacity Building:* Evaluations can provide an opportunity to better engage stakeholders in the program or process. Involvement can increase stakeholders' knowledge of the program and make them feel they have a voice. Stakeholder ownership can stem from participation at various stages and levels, from the more common role of being used as a resource during data collection to more in-depth functions such as generating evaluation questions, developing the evaluation design, or collecting data.

6) *Strategic:* Credible evaluations can persuade others or gain particular strategic outcomes.

Evaluation Requirements

All bureaus and independent offices with a budget of one million dollars or more are required to conduct at least one evaluation each year. Larger offices are expected to evaluate more of their programs, projects, and processes, commensurate with the scope of their portfolio, size of their budget, and perceived needs of management. Large programs are defined as those falling above the median for any bureau or office with reference to the budget allocated or full-time equivalent staff dedicated to them.

Pilot programs or interventions, defined as any new, untested approach that is anticipated to be expanded in scale or scope, must be evaluated for impact before being replicated or expanded. Pilot interventions should be identified during program design, and the impact evaluation should be integrated into the design of the project or activity. If it is not feasible to effectively undertake an impact evaluation, bureaus and independent offices must conduct a performance evaluation and document why an impact evaluation was not feasible. In some cases, small, relatively low cost interventions will not rise to the level of a pilot, but fall more in the realm of proving the viability of a concept. While these pre-pilot interventions may not warrant the costs and effort of an impact evaluation, bureaus should test whether the intervention improves the baseline. This can be done through a variety of methods such as pre and post assessment, before and after comparisons, and testing versions of an intervention for

comparison. These proof-of-concept interventions may or may not eventually become pilots. Bureaus and independent offices should use professional judgment in designating programs as pilots during the planning and design phase.

Bureaus managing multi-year programs that continue indefinitely should conduct comprehensive evaluations to examine their performance and impacts at least once every five years. If indefinite programs are supported by contracts, evaluations should be timed to inform subsequent contracting decisions.

Bureaus that administer presidential initiatives involving numerous long-term programs and projects must ensure that evaluations are completed. These initiatives should also develop formal guidance and procedures for evaluations, including building it into program planning and ensuring dissemination of report results. Bureaus that make voluntary contributions to international organizations should ensure the organizations have evaluation standards and procedures aligned with State's policies. In addition, bureaus should ensure that U.S. government agencies to whom they pass funds for foreign assistance have effective monitoring and evaluation standards in place and report on the projects or activities the money funded. This helps ensure the Department's accountability for its funds. For detailed information, See Transfers of Funds in Section 8 of this document, Implementation.

Bureau Evaluation Coordinators

According to 18 FAM 300, each bureau must appoint a "Bureau Evaluation Coordinator" (BEC) with decision-making authority to promote evaluations in bureaus and independent offices. BECs should have expertise or training in evaluation so they can perform the following functions:

- *Bureau Evaluation Plan:* The BEC should work with bureau leaders and program managers to put in place an evaluation plan that is tied to the bureau and the Department's strategic priorities.
- *Promote Use of Evaluation Results:* The BEC should convene meetings of program managers to facilitate the utilization of the findings and recommendations of evaluations.
- *Ensure Bureau Senior Leadership Engagement:* The BEC should brief senior leaders and other staff after an evaluation is completed.
- *Share Lessons Learned:* The BEC should share lessons learned from past evaluations to aid the planning process.
- *Facilitate Corrective Actions:* The BEC should work with program staff and subject-matter experts to ensure corrective actions, identified during the evaluation, are satisfactorily completed.

Bureau Evaluation Plans

Bureaus and independent offices must complete a bureau evaluation plan. The plan should cover prior-year completed evaluations, ongoing evaluations, and those that are planned in the next one to two years. These evaluations should be reflected in the BRR or MRR processes. Foreign assistance-funded evaluations are recorded in the Evaluation Registry while diplomatic engagement funded evaluations are recorded in the Evaluation Management System. The compilation of the two systems constitutes the bureau's evaluation plan. Bureaus and independent offices with large budgets or many programs should also consider creating a separate plan that provides more detail and analysis of what should be evaluated. They may refer to the [Bureau Evaluation Plan](#) templates found on the Resources page of the [Evaluation Community of Practice site](#) (under Bureau Status of Evaluation Reporting Requirements) for options on creating more detailed plans outside of the Registry or EMS.

Considerations for Evaluation

While it is up to a bureau or independent office to decide what programs, management processes, projects, services, initiatives, activities, and delivery systems should be evaluated, bureaus and independent offices may take into account the following considerations in making the decision. Templates for three different approaches to planning for evaluation are found on the Resources page of the Evaluation CoP under [Bureau Status of Evaluation Reporting Requirements](#).

- **The priorities of the bureau or independent office:** Certain efforts are major priorities and are reflected in Functional Bureau Strategies, Joint Regional Strategies, or the State-USAID Joint Strategic Plan. They include foreign assistance interventions, management, diplomatic initiatives, and efforts the Administration, the Congress, or the Department have identified as being of major importance. Depending on the number of goals in their FBS or JRS, bureaus and independent offices should attempt to ensure that at least an aspect of each goal of the strategy is evaluated over the life of the strategy.
- **Information needs:** Bureaus and independent offices should examine their own information needs about the performance and outcomes of ongoing projects, programs, or processes. Bureaus may give preference to the evaluations of those activities about which decisions are to be undertaken regarding their continuation, modification, or termination.
- **Innovativeness of the program or projects:** Sometimes new innovative programs and projects are designed and implemented by bureaus and independent offices. These may be pilots, which if successful may be replicated elsewhere. Please refer back to the definition of pilot and relevant evaluation requirements on p. 11.

- **Program, project, or process size:** As required, bureaus and independent offices should also consider the size and complexity of a program, project, or process. Large programs, projects, or processes should get precedence over smaller ones. When all are essentially of the same size and type, bureaus and independent offices should consider the other criteria listed here in order to fulfill the requirement.
- **Feasibility of data collection:** Bureaus and independent offices should consider whether relevant data and information for an evaluation can be gathered within the given time and resources. In many cases, even when information is needed, evaluation teams may not be able to gain access to it. It is often difficult to gather reliable and valid data in high threat environments or in cases when the beneficiaries have long since dispersed.
- **Opportunity to build bureau capacity:** Some evaluations may lend themselves to increasing knowledge in the bureau or independent office about program planning, monitoring, or execution. For instance, sometimes a program needs an evaluation but has no systematically collected data. A logical framework or a monitoring plan can be added to the statement of work (SOW) as a deliverable. A bureau may also know that the project or program may have difficulty implementing changes or recommendations. In this case it can ask for an implementation road map or a change management guide as part of the evaluation.

Ethical Standards

Department staff and evaluators should follow the accepted ethical standards in dealing with stakeholders, beneficiaries, and other informants. These are

- **Rights of Human Subjects:** When human subjects are involved, evaluations should be designed and conducted to respect and protect their rights and welfare in accordance with Title 22 of the US Code.
- **Sensitiveness:** Evaluators should be sensitive to the gender, beliefs, manners, and customs of people and conduct their research in a culturally appropriate fashion.
- **Privacy and Confidentiality of Information:** The privacy and confidentiality of information should be maintained. If sensitive information is involved, efforts should be made to ensure that the identity of informants is not disclosed in accordance with Title 22 of the US Code.
- **Conflict of Interest:** Care should be taken that evaluators have no potential biases or vested interest in the evaluation outcomes. For example, a firm should not be

contracted to conduct an evaluation if they played any role in supporting the execution of that program or activity being evaluated.

Types of Evaluation

Evaluations are designed to determine relevance, effectiveness, efficiency, sustainability, or impact. As such, they should:

- Focus on current or completed programs, projects, and management processes. Thus planning documents (including assessments), appraisal reports, and exploratory studies should not be treated as evaluations.
- Examine one or more of the following: effectiveness, efficiency, sustainability, continued relevance, or impact of a project/program or management activity.
- Be based on systematic collection and analysis of data and information by experts who are not accountable to managers of the project, program, or activity or process. Therefore progress, monitoring, or inspection reports submitted by a program or implementing agency and field-trip reports undertaken by program managers are not evaluations.

With these criteria in mind, and depending upon their information needs, resources, and priorities, bureaus can undertake a wide range of evaluations, which may include but are not limited to the following:

Performance evaluations focus on the performance of a project, program, diplomatic activity, or process and examine its implementation, inputs, outputs, and likely outcomes. Performance evaluations of ongoing programs, projects, and processes usually evaluate their efficiency, effectiveness, and relevance. They may reveal whether the programs were effectively managed or provided planned goods and services in a timely fashion, as well as whether they met targets and were cost effective.

Performance evaluations focus on questions such as: Were the targets or objectives met? Are the current procedures and practices the best way to achieve the intended results? Are there other alternatives which can be more effective and efficient? For example, the Bureau of Consular Affairs may want to evaluate if its system for reviewing visa applications online is functioning well. Is it cost effective? Will use of more sophisticated technology reduce time processing visa applications? An overwhelming majority of evaluations conducted by bureaus are performance evaluations.

There is no hard or fast rule regarding when bureaus should conduct performance evaluations; however, for a project/program with a life span of five years, bureaus should consider undertaking an evaluation during its mid-cycle so that managers can have an objective assessment of its implementation progress, problems, and challenges, which would enable them to make mid-course corrections, if necessary. On the other hand, for a project/program with a duration of two or three years, it might be preferable to conduct the evaluation at its end because it can take eight to twelve months before a project/program becomes operational and its outputs are visible.

Process evaluations examine the quality or efficiency of program operations, determine whether a program was executed as planned, and consider the problems and challenges the program faced. They focus on how well activities are carried out and whether they reach intended recipients.

Organizational evaluations (also known as organizational assessments) examine business operations, policies, procedures, and personnel resources. They can range in scope from analyzing unit or division activities to a bureau-wide analysis of operations. Such evaluations usually make recommendations to help bureaus to execute their missions more efficiently and effectively. For example, the Bureau of African Affairs sponsored an organizational evaluation to inform its strategy to effectively support long-range planning and strategies.

Sector evaluations examine the performance and outcomes of major projects and programs in a sector or sub-sector to develop general findings, conclusions, and recommendations. A global evaluation of electoral assistance, for example, will not evaluate the success or failure of individual projects or programs but will determine the efficacy and outcomes of entire electoral assistance programs and practices as a whole. The country level sector evaluations are particularly helpful in deciding about future programming. For example, INL may undertake a major evaluation of all its security sector projects in a war-torn society to assess their impacts and to make recommendations about what types of foreign assistance interventions should be undertaken in the future. Bureaus may conduct sector evaluations at country, regional, or global levels.

Summative/ex-post evaluations differ from performance evaluations in that their focus is primarily on outcomes and impacts but often include effectiveness, and they are conducted when an effort has ended or is soon to end. Summative/ex-post evaluations answer questions such as: What did the program achieve? What changes were observed in targeted populations, organizations or policies during and at the end of effort? To what extent can the observed changes be considered to have contributed to it? Were there unintended effects which were not anticipated at the planning stage? Were they positive or negative? What factors explain the intended and unintended effects? The essence of summative/ex-post evaluations lies in

establishing that the intervention has substantially contributed to the changes that have occurred.

Impact evaluations measure the change in an outcome that is attributable to a defined intervention; they are based on models of cause and effect and require a credible and rigorously defined counterfactual (comparison group) to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between like populations 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. The overall effects are measured by comparing the performance, conditions, or status of the two groups. Establishing control groups to generate counterfactual data must be done at the beginning with comparable groups. Performing an impact evaluation is not always feasible or practical because of the advance planning, cost, or the inability to create a control group.

A **meta-evaluation** involves reviewing completed evaluations against a set of professional standards. A meta-evaluation may examine evaluation design and data collection strategies, evaluators' interpretations of the data, the links between findings and recommendations and even how the findings are being used. There are no set standards for meta-evaluations, and different evaluators tend to use different criteria. Evaluators in a wide range of fields—education, health, climate, family planning, and enterprise development—have used varying criteria depending upon their needs and disciplinary orientations.

Collaboration and Other Ways to Evaluate

Bureaus and independent offices may conduct internal, external, hybrid, or collaborative evaluations, described below.

Internal Evaluations: When a bureau conducts evaluations by its own staff, these are known as internal evaluations. Many bureaus have their own qualified evaluators to design and implement their evaluations. Internal evaluations, a type of low cost/no cost evaluation, have many strengths. Since evaluators are embedded in bureaus, they possess an intimate understanding of the political and social context of a bureau or office's programs, management processes, and delivery systems. Internal evaluators are likely to be invested in providing useful and actionable findings, increasing the likelihood that findings will be used. They are also acquainted with program managers and therefore understand their needs, concerns and expectations. Moreover, internal evaluations can take less time, as bureaus need not contract them out. There are two essential requirements for internal evaluations: (a) the bureau has trained evaluation staff, and (b) the evaluation staff is not in a direct supervisory line to the managers of the program or the processes to be evaluated.

Internal evaluations can be done with little or no cost with TDY personnel or staff within a bureau not directly responsible for the activity or program being evaluated. Such evaluations typically focus on a limited number of questions relating to the performance or limited outcomes of a project. They use rapid, low cost data collection methods. A bureau or independent office may use them for evaluating a small project, program, or activity where the cost of a full, external evaluation may come close to the overall cost of the project itself.

F and BP do not require bureaus to develop a formal SOW for an internal evaluation. However, internal evaluations should proceed from a framework that reflects a high degree of rigor. While bureaus and independent offices have the option to develop a SOW, they may also develop a short document that includes the following components:

- Title of Evaluation
- Evaluation Objective(s)
- Bureau/Office Conducting the Evaluation
- Evaluation Questions
- Proposed Methodology

The briefer documents should be treated as SOWs for purposes of reporting in the Evaluation Registry or the Evaluation Management System.

External Evaluations: Since the majority of Department of State bureaus and independent offices do not have professional evaluators on their staff, many evaluations are likely to be contracted to outside firms and organizations. Although more expensive and time consuming, external evaluations have three advantages over internal evaluations. First, since evaluators are outsiders, they enjoy relatively more freedom than embedded evaluators and are usually less susceptible to the influence of program managers. Second, outside firms and research organizations typically have access to a large pool of highly trained evaluation professionals, who possess technical expertise and experience. Third, some decision-makers give greater credence to the findings and recommendations of external evaluators than to evaluations performed by their own staff.

Hybrid Evaluations: Hybrid evaluations involve both internal and external staff working together and try to bring together the advantages of both internal and external evaluations. Internal evaluators bring the benefit of being more familiar with a program, project, or process while external evaluators bring specific expertise from which a program might benefit and are often perceived to be more independent. There are different ways to organize a hybrid evaluation, including having external and internal evaluators work alongside each other on the evaluation team or the external evaluator supports internal staff to conduct an evaluation through facilitation and/or coaching, and just-in-time technical advice. A hybrid evaluation may

be most relevant when specific technical or cultural competence is needed that requires a mix of evaluators, or when the credibility of findings will be enhanced by having a mix of internal and external staffing.

Collaborative Evaluations: The essential idea behind collaborative evaluations is to pool resources to conduct an evaluation. All collaborators jointly develop a SOW and engage in the entire evaluation process. The nature and structure of their participation may vary, depending on the time, resources, and expertise available to them. Although such evaluations are usually conducted of a single project, program, or process, sometimes complementary projects, programs, or processes may also be evaluated collaboratively. In such cases, evaluators do not examine them separately but as a collective endeavor, and write a report that synthesizes the findings, conclusions, and recommendations of all projects or programs evaluated.

Collaborators may include bureaus, agencies, universities, or even international partners. For example, regional bureaus may wish to collaborate with functional bureaus to evaluate a project or program. Collaboration can also involve other agencies or international partners. For an example of a multi-agency, international evaluation, the Department participated in an international evaluation of the implementation of the Paris Declaration on international assistance. Ample opportunities also exist for bureaus and independent offices to collaborate with their counterparts in USAID, DOD, or other agencies to evaluate jointly funded programs.

Collaborative evaluations have many advantages. They facilitate mutual learning among participating organizations. The burden of conducting evaluations is shared among the participating organizations, thereby reducing the cost incurred by all. Finally, as a result of pooling resources, more intensive and systematic data collection and analysis can be undertaken, which enhances both the quality and credibility of the evaluation. On the negative side, collaborative evaluations tend to be time consuming and require strong leadership to reconcile conflicting interests, priorities, and expectations of collaborators. Each participating bureau will be able to report the evaluation in their evaluation plan under the Design, Monitoring, and Evaluation Policy as long as they have a written agreement or an MOU outlining the roles and responsibilities of each party.

Participatory Evaluations: Such evaluations involve the participation of multiple stakeholders in the evaluation process, which includes setting evaluation objectives, defining priorities, selecting questions, analyzing and interpreting the data, and framing recommendations. For example, the evaluation of a public diplomacy project, which awards scholarships to journalists for visiting the United States, would require that the bureau, the implementing organizations, the stakeholder embassies, and journalists would jointly undertake the evaluation with or without the assistance of technical experts. The main strength of participatory evaluations is that the managers listen and respond to the views and recommendations of stakeholders

during the evaluation process. Face to face interactions also facilitate a better understanding of the working of a process or intervention. However, the scope of such evaluations is bound to remain limited in the Department because all stakeholders are generally not located in Washington or in a post. The distinction between collaborative and participatory evaluations is that the Department in most cases assumes the cost burden for the participatory evaluation.

Planning and Managing Evaluations

It is essential that bureaus and independent offices have comprehensive plans for all their programs and projects. The plans should provide details about the overall objectives, underlying logic model, expected outcomes, outputs and targets, time table for implementation, and implementation schedule. Project plans should also have budgets for monitoring and evaluation built into the original costs.

Bureaus and independent offices are strongly encouraged to develop effective performance management systems that regularly monitor activities, projects, and programs based on carefully selected indicators. The indicators could be both qualitative and quantitative; however, it is necessary that data are regularly collected and analyzed. Such performance systems are essential for sound management. Moreover, they provide data and information for conducting rigorous evaluations. For more information about design and monitoring, please refer to sections 4 and 5 of this guidance.

Specifying Objectives and Audience

Specification of objectives helps in sharpening the focus of an evaluation and in formulating evaluation questions. It also makes it easier for evaluators to come up with relevant findings and recommendations. An evaluation's objectives should be stated in such a way that they also specify how the generated information will be utilized. For example, it is not enough to state that the objective of an evaluation is to examine the performance of an electoral assistance project or assess its impacts. It is also necessary to specify, whenever possible, how the findings will be used to improve the performance of the project or how the lessons of summative evaluation will feed into planning new projects in that area. To this point, bureaus should create dissemination plans for evaluations to ensure that the evaluation reaches the broadest audience and its objectives are met. A template for an evaluation dissemination plan is on the [Resources](#) page of the CoP website.

In addition to considering their own information needs, bureaus should also consider the needs of other stakeholders – implementing organizations, the host country, or other bureaus in Washington – who might be interested in the evaluation and could benefit from it.

Formulating Evaluation Questions

Questions should grow from the objectives of the evaluation. An evaluation cannot answer too many questions effectively, so spending the necessary time to focus on what questions will best

answer information needs is important. Most evaluations should have no more than five questions. We recommend using the Design Matrix, available on the [Resources page](#) of the Evaluation Community of Practice site, to formulate questions. Filling out the matrix will help evaluation planners think through not only what they want to ask, but whether the questions are answerable with the time and resources available. For instance, the matrix asks what kind of information will be needed to answer the questions and what methods will be best for collecting it.

Determining Data Collection Methods

Evaluation design and data collection methodologies should be appropriate to answer the key questions posed by the evaluation, and the data can be both qualitative and quantitative. Data collection methods may include:

Document Reviews: A systematic review of relevant literature is invaluable for evaluations, and provides essential information about the nature, scope, achievements, and challenges facing the effort to be evaluated. It aids in analyzing the intervention model that underpins the effort. Above all, it gives tentative answers to the questions the evaluator would further explore in fieldwork.

Secondary Analysis of Data: There are numerous national and international organizations that routinely collect data on a variety of variables relevant to Department activities. Often they are time series data covering a span of years, which is an added advantage. For example, rather than conduct a survey about peoples' perceptions of democracy in a country recovering from a prolonged civil war, evaluators might access existing surveys of a research organization.

Key Informant Interviews: Key informant interviews involve interviewing a select group of individuals for information, ideas, and insights on a particular topic. Two characteristics of key informant interviews need special mention. First, only a small number of informants are generally interviewed, who are selected because they possess information or ideas that can be useful to evaluators. Second, key informant interviews are qualitative interviews that involve continually probing the informant. The atmosphere in these interviews is informal, resembling a conversation among new acquaintances. Virtually all evaluations use key informant interviews. Even those based on quantitative data use them for constructing research instruments and interpreting the gathered data.

Surveys: Surveys are extremely useful in gathering quantitative data. Well-designed surveys can generate a wealth of data for evaluations of programs and can improve both their quality and relevance. In surveys, only a fraction (sample) of the entire population is selected using probability sampling. Surveys use a structured questionnaire that is administered to respondents. Depending upon the scope of the evaluation, respondents may provide

information about themselves, others, or the organizations they know. It must be noted that the development of a valid sample is highly technical and requires appropriate professional expertise.

Mini Surveys: Mini-surveys differ from surveys in three respects. First, they focus on a few questions, usually 10 to 25. Second, the sample size is smaller. Mini-surveys employ non-probability sampling techniques to select respondents. In most cases, a set of criteria is established, and a limited number is selected on the basis of their availability. Third, analysis is limited to descriptive statistics. Mini-surveys are particularly useful when the evaluation team has limited time for fieldwork because they can be mounted rapidly and provide some quantitative data that can complement information gathered from other sources. Because they do not use probability sampling, caution is necessary in generalizing from their basis.

Structured Direct Observation: Another data collection method that can generate reliable, relevant, and quantitative data is structured direct observation, which involves observing interactions, processes, and behaviors as they occur on the basis of structured observation forms. This method can be used for gathering information about programs such as election monitoring, on-going training, or delivery of health services. The main strength of direct observation is that an event, institution, facility, or process can be studied in its natural setting, thereby providing a rich understanding of the subject.

Social Network Analysis: Though not strictly a data collection method, social network analysis (SNA) has emerged as a powerful tool in the arsenal of evaluators. SNA involves mapping and measuring relationships between individuals or groups. The unit of analysis is not an individual but the network of actors. The underlying assumption is that actors and their actions in a network are interdependent and influence each other although in varying degrees. The linkages between them are the channels through which information, ideas, and resources flow. The most common way to collect necessary data is to ask individuals in a network to identify their relationships with each other. Usually a questionnaire is administered for this purpose. Evaluators have widely used SNA in evaluating a wide range of development projects and programs in health, rural development, and community development sectors. They have also used it to conduct evaluations of democracy promotion and security sector interventions. Software that has simplified the task of coding and analyzing network data is now freely available.

Focus Group Discussions: A focus group discussion is essentially a planned dialogue among a small group of people in a naturalistic environment. The purpose of a focus group discussion is to solicit information and ideas from participants. Focus group discussions differ from small group interviews in that they use interaction among group members to elicit information. It involves free and spontaneous discussion on a set of specific topics in which six to 12 people

participate with the aid of a moderator. The moderator's job is to stay largely in the background, primarily facilitating the discussion within set boundaries. Some also expect focus group members to possess some degree of homogeneity, which may set it apart from a small group interview. When conducted properly, focus groups are a powerful instrument which can provide an in-depth understanding of a phenomenon, process, or event.

Selecting an Evaluation Design

The questions, resource needs, data collection methods, timeline, data availability, and limitations thought through in the Design Matrix will help planners select an appropriate design. In some cases, the previous planning for a program or project may also dictate the method, depending on the monitoring and reporting done. In many cases an experience review or desk review can serve as a precursor to a final design. These reviews involve systematic analysis of past experience of projects, programs, or management processes mostly based on project/program documents, monitoring data, and evaluations. They can generally be completed within a month depending upon the scope of work and the volume of documents and reports to be reviewed. Information gleaned can help refine the evaluation questions in your design. Evaluation designs most often used at the Department are:

Mixed Method Design: The most common design is the mixed method design. In such evaluations, the evaluation team uses both qualitative and quantitative data to answer evaluation questions. Each evaluation question is answered by one or more methods. When two methods are used to answer the same question, both types of data are presented and their findings synthesized. Such evaluations enable triangulation of data and information, which improves the reliability and credibility of the evaluation findings. An overwhelming majority of the evaluations conducted by the Department use this design.

Case Study Designs: The case study design examines a program, management process, or intervention in its context, focusing on the dynamic interactions between it, its environment, and the concerned actors. A case study often uses a combination of quantitative and qualitative data. Evaluators choose particular and limited instances of the evaluated subject to represent the changes in the variables of interest. Case studies explore issues in depth as evaluators enjoy relatively more freedom in pursuing clues generated during field research as compared to other designs. In addition to document reviews, they rely on direct observation of the sites, interviews with key informants, and focus group discussions. In some cases, they may even conduct informal surveys. Evaluators can also construct case studies of different interventions, programs, or management processes and synthesize their findings.

Quantitative Designs without Control Groups: Such designs primarily use data derived from sample surveys, content analyses of documents, and structured observations to answer evaluation questions. There are many variants of quantitative designs. One consists of pre and

post design, which measures performance and impacts before and after an intervention is over. For example, surveys are undertaken in a narcotic crop producing area before and at the end of a program. The decline in narcotic crop production, if any, can then be attributed to the program. In another variant of this design, time series data are collected at several points in time in order to account for fluctuations. Still another category includes cross-sectional analyses in which outcomes are measured by analyzing varying exposures to project's activities that participants received. The advantage of such designs is that they generate precise statistical data which key decision-makers and policy-makers often find credible. Moreover, in many cases, they can measure impacts without counterfactuals. The opportunity for using rigorous quantitative designs for Department activities remains limited.

Experimental and Quasi-Experimental Designs: In experimental and quasi-experimental designs, evaluators compare the results of two groups – one that received treatment and the other (control) that did not. The impact is measured by comparing the treatment and non-treatment groups before and after the completion of the program or project. Experimental and quasi-experimental designs require quantification of the outcomes and the construction of control groups, which includes finding and measuring comparable problems in similar populations in comparable environments who are not in contact with each other. They also require that once the program or project is launched and the baseline data is collected, there are no changes in how the program or project operates during its implementation. These designs require that the information, ideas, innovations, or material goods provided to the treatment group are not available to the control group during the life of the program or project. Above all, they usually require heavy investments in data collection.

Preparing a Statement of Work

A well-written and well-thought-out SOW⁴ ensures that the evaluation will meet the information needs of the sponsoring bureau or independent office. It provides guidelines to evaluators and reduces possible misunderstandings that may arise between them and evaluation managers. Even when an internal evaluation will be conducted, a SOW may be valuable for analytical and management purposes.

A SOW is essentially a blueprint for an evaluation which provides a road map to evaluators. It gives details about the proposed evaluation, its focus and objectives, the questions it is supposed to answer, appropriate evaluation design and the data collection methods it will use, the expertise of the evaluators and composition of evaluation teams, and the deliverables it will

⁴ Occasionally, a Performance Work Statement (PWS) is used to solicit for evaluation expertise for management processes. A PWS is synonymous with statement of work.

produce. It also includes a budget and a time table. A SOW is required for awarding contracts to outside firms or research organizations.

Preparing a SOW requires reflection, consultation, and preliminary research. Although the time and effort required varies depending upon the nature, complexity, and purpose of an evaluation, preparing a good SOW takes time. If the evaluation manager lacks experience in preparing them, s/he should consult colleagues in their bureau, F and BP, and even outside experts who will not be engaged in the conduct of the evaluation. The manager may also share the first draft with colleagues to solicit comments and suggestions.

F and BP staff can review draft SOWs to help ensure that they are complete and accurate. Please allow sufficient time for an in-depth review; BP staff will review SOWs of diplomatic engagement-funded evaluations and F staff will review SOWs of foreign assistance-funded evaluations. (Note that bureaus using the IDIQ must submit draft SOWs to BP or F (according to funding being used) for review. Visit the [contracting page on the Evaluation CoP](#) site for more information.)

An evaluation SOW should cover the following:

- Purpose of the evaluation – scope, audience, and intended use of findings
- A brief history and current status of the project/program/strategy to be evaluated
- Evaluation questions
- Available information and sources – performance data, FACTS data, previous evaluations
- Preliminary evaluation design and data collection methods
- Qualifications of evaluators and composition of evaluation team
- Deliverables (e.g., work plan for the evaluation, draft and final evaluation reports, briefing with stakeholders, data placemats, executive summaries, etc.
- Time schedule (period of performance)
- Budget (Independent Government Cost Estimate, or IGCE)

The SOW guidance package includes step by step instructions and templates. It is available on the [Evaluation Community of Practice](#) page under [Resources](#).

Contracting Evaluations

F and BP have an umbrella IDIQ evaluation contract with selected evaluation firms. If a bureau chooses to use the mechanism, firms with a contract under the IDIQ have the opportunity to compete by submitting an evaluation proposal, and the bureau would choose the best proposal through a technical evaluation panel process. Bureaus are not required to utilize the IDIQ firms. There are other mechanisms that may be used to contract evaluations.

In addition, the Department has Blanket Purchase Agreements (BPA) to facilitate evaluation-related services such as evaluations, process and performance improvement, strategic planning, performance management, and business analysis. Bureaus and independent offices are encouraged to explore these State-approved contracting vehicles when assessing whether to use the umbrella IDIQ or an alternate contracting mechanism. Information on [contracting vehicles](#) is available on the [Evaluation Community of Practice](#) page.

Criteria for Selecting Evaluators Proposed by Contractors

Bureaus and independent offices should ensure that the evaluation teams proposed by contractors possess professional training and experience in conducting evaluations. Teams should be able to develop appropriate evaluation designs, collect data and information using various data gathering methods, establish rapport with stakeholders, and write empirically grounded reports with actionable recommendations. Moreover, they should have substantive knowledge of the subject to be evaluated.

In overseas evaluations, it is often advisable to include host-country evaluators. Because they provide insiders' perspectives, speak local languages, and even have access to studies and documents that are not always available to outside evaluators, their participation tends to improve the analytical rigor of evaluation.

Gender representation in the evaluation team usually aids in data collection. For example, women evaluators are often more effective in soliciting information and ideas from women respondents, who tend to be more comfortable with them than with male interviewers. The reverse is also true in many traditional societies where men are reluctant to be interviewed by women. When the target population is mixed, bureaus and independent offices should insist that the contracting firm pay particular attention to the gender mix of evaluation teams so that they are able to conduct their research with all stakeholders in a culturally appropriate manner.

Preliminary Discussions with Contractors after Award

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

- Evaluation questions: Do questions need refinement? Should additional questions be added and/or existing ones deleted?
- Evaluation design: Managers and evaluators should explore options other than those proposed in the SOW, if necessary. Evaluators should submit a draft evaluation design

for review by the manager. Managers should ensure that the draft design is comprehensive and reasonable. This includes, for example, checking that methods are elucidated by and appropriate to each question, limitations are clearly stated, sampling methods are detailed and logical, and any proposed changes to the evaluation questions align with the original evaluation purpose.

- Time and resources: Are the amount of time and level of effort provided sufficient?
- Nature and content of evaluation report(s), briefs, and presentations

In light of these discussions, the evaluation team should develop – or update, if one was submitted during the proposal process – a comprehensive evaluation work plan (possibly through a team planning meeting). At its most basic, a work plan should include who is doing what and when. Upon receipt of the draft work plan, managers should ensure all deliverables are listed, staff proposed in the plan align with those proposed in the original proposal, timelines presented are reasonable and sufficient, level of effort is within the level agreed upon, travel dates do not coincide with holidays or other times that make travel challenging, and there is a deadline for confirming an interview schedule prior to departure for fieldwork.

Evaluation Report Review and Follow up

Bureaus and independent offices should insist that evaluation reports be clear, concise, empirically grounded and readable. As much as possible, reports should strike a reasonable balance between depth and length. When reports are lengthy, readers may be inclined to put them off or only casually read them. On the other hand, if reports are too brief, vital information cannot be presented. The best approach is to have a report of between 25 to 35 pages, excluding annexes. The report should use simple and straightforward language and avoid obscure expressions and technical jargon.

In addition to an executive summary, most reports should have five components. The first component should describe the activity being evaluated, the purpose of the evaluation and evaluation questions. The second should include a succinct description of the evaluation design and data collections methods and their limitations. The third should present data and findings coherently and be organized by evaluation question. The fourth component should be the conclusions, which represent the evaluators' judgments based on the findings. The fifth, and in many cases, most critical element of an evaluation report should be recommendations, which should flow logically from the findings and conclusions. They should be practical in the sense that the concerned decision makers have the authority and resources to implement them. Recommendations should not only mention what needs to be done but how it should be done. When recommendations are directed to multiple stakeholders, they should specify which recommendation should be implemented by which stakeholder. Details about research methods, sampling, or research instruments should be included as annexes so that those who want to go into more depth can read them. Further information on the organization and contents of an evaluation report is found on the [Resources](#) page of the CoP website.

Technical Assistance and Capacity Building

The evaluation teams in both F and BP can provide technical assistance to interested bureaus and independent offices in planning and conducting evaluations. This includes

- Helping bureaus and independent offices plan evaluations and develop questions.
- Reviewing SOWs prior to contracting.
- Providing general guidance for reviewing evaluation reports.
- Working with bureaus to ensure compliance with evaluation procedures and requirements.
- Facilitating formal training of Department staff and contractors in monitoring and evaluation through the State-sponsored courses “Managing Evaluations” and “Evaluation Designs and Methods”.
- Organizing workshops, meetings, lectures, and conferences to promote the institutionalization of evaluation in the Department.
- Maintaining a website to facilitate information sharing and learning (see [Evaluation Community of Practice](#)).
- Preparing documents, manuals, and articles on evaluations for the benefit of evaluation managers (see [Evaluation Community of Practice](#)).
- Facilitating and promoting the utilization of evaluation findings.
- Representing the Department in national and international forums and networks on evaluation.
- Selecting the appropriate procurement mechanism
- Facilitating use of the Performance Management and Evaluation Services IDIQ
- Facilitating learning and the exchange of information through the [Evaluation Community of Practice](#).

Evaluation Use

The information, ideas, and recommendations generated by evaluations should serve specific needs of the Department. More specifically, they should serve one or more of the following objectives: (a) improving ongoing Department efforts by examining their implementation and by generating actionable recommendations, (b) improving future operations by codifying experiences and lessons learned (what works or does not work), accompanying corrective actions, and (c) enhancing the Department’s accountability to major stakeholders – the White House, OMB, the Congress, taxpayers, and host countries – by systematically examining if existing efforts are achieving what they were intended to achieve.

Bureaus and independent offices conducting evaluations should plan for the evaluation’s use early on by preparing a dissemination plan. The dissemination plan should include at a minimum a list and description of the stakeholders who will benefit from and receive the evaluation results, methods for reaching those stakeholders, and an estimated budget. As a

dissemination plan is a key part of learning activities that accompany an evaluation, budgeting for evaluation can include the design of the report, development of presentations, customs graphics, or any other product that will assist stakeholders in understanding the results and acting on recommendations. Please see section 8 of the Guidance for additional information on budgeting for evaluations.

Once they have a report, bureaus and independent offices should prepare a written summary of the recommendations for bureau or office leadership. This will allow management to discuss the recommendations, determine whether they concur, create a plan for implementation within a certain timeframe, and designate a point of contact for implementation. Bureaus and independent offices should monitor progress on follow-up to the recommendations through a document such as a recommendation tracker. The tracking document should be used until recommendations are implemented. The CoP [Resources](#) page has templates showing three approaches to recommendation follow-up. Bureaus and independent offices should use evaluation findings to make decisions about policies, strategies, program priorities, and delivery of services, as well as for 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, or to shape that plan initially.

Dissemination

Dissemination is handled differently depending on whether the evaluation was funded by foreign assistance or diplomatic engagement monies.

Foreign Assistance

FATAA requires that completed reports of all evaluations of foreign assistance projects or programs be made available on a public website and be disseminated internally within agencies for learning. F published *Guidance for Public Posting of Foreign Assistance Evaluations and Summaries* in 2014. Available under [Resources on the CoP](#) page, the guidance stipulates that the default for reports of foreign assistance evaluations should be to post publicly. For foreign assistance evaluation reports with sensitivities, a publishable summary is required and should be indicated in the SOW as a separate deliverable. Classified evaluation reports are exempt from the public dissemination requirement entirely. Evaluations determined to be Sensitive But Unclassified must publish a summary of results, but are not required to publish the entire report. Reports should be compliant with Section 508 of the Rehabilitation Act of 1973, as amended in 1998 (29 U.S.C. § 794 (d)), which provides accessibility to electronic information in concert with the Americans with Disabilities Act. The Department recommends that bureaus write the requirement for Section 508 Compliance into SOWs.

F collects reports and summaries of evaluations funded by foreign assistance in the Evaluation Registry. The reports and summaries, if designated for public use, are then posted on the public

[Foreign Assistance Evaluations](#) page. Reports or summaries, as appropriate per guidance, of reports for foreign assistance evaluations must be made public within 90 days of their issuance.

Bureaus and independent offices that administer very large foreign assistance programs or special initiatives through contracting for numerous long-term programs and projects, or which jointly fund or participate in evaluations through international organizations, should ensure copies of 508 compliant summaries or full reports are provided to F for public posting or that they are available on a public website F can link to from state.gov.

Diplomatic Engagement

All evaluations funded by diplomatic engagement funds remain internal documents that, per the Evaluation Policy, are made available internally to all State Department bureaus and offices for discussion and learning via the [Evaluation Management System](#). Evaluations that may be sensitive or confidential for a bureau or office can be posted to the EMS and tagged appropriately to limit availability. At their own discretion, bureaus may elect to make diplomatic engagement-funded evaluation reports available to a wider audience though it is not a requirement under this policy.

7. Analysis and Learning

Senior Department bureau leaders and chiefs of mission will institute regular reviews to assess progress against strategic objectives, and ensure alignment of policy, planning, resources, and program decision making. To implement this requirement, BP and F have developed strategic progress review [guidance](#) to assist bureaus and missions. Additionally, guidelines and templates in the [Program Design and Performance Management Toolkit](#) may assist bureaus and missions in preparing for and executing their reviews. Complete guidance, supplemental materials, the PD/PM Toolkit, and sample templates can be found on the Managing for Results (MfR) website: <http://cas.state.gov/managingforresults/>.

8. Implementation

F and BP oversee implementation of the 18 FAM 300 and can provide technical assistance for program or project design, monitoring, evaluation, and learning activities. F assists with foreign assistance-funded activities and BP assists with those funded by Diplomatic Engagement.

BP oversees the [Evaluation Management System](#) (EMS), which is a web-based, customized system that serves as the system of record for all evaluations funded by diplomatic engagement funds and for all Bureau Evaluation Plans for bureaus operating with diplomatic engagement funds. The EMS captures information on planned, ongoing and completed diplomatic engagement-funded evaluations, including reports and information on intended and actual use of evaluation findings. An annual data call instructs bureaus and independent offices to update

and/or input new evaluation records in the EMS, resulting in the establishment of Bureau of Evaluation Plans for diplomatic engagement evaluations.

F manages the Evaluation Registry, the system of record for all evaluations funded by foreign assistance. Housed in [FACTS Info NextGen](#), the Registry serves as a reporting and management tool for bureaus and independent offices to record and track planned, ongoing, and completed evaluations. The combination of records for a bureau or independent office in any one fiscal year serves as its Bureau Evaluation Plan for its foreign assistance-funded evaluations for that year; no additional document is required. Bureaus and independent offices reporting evaluations of foreign assistance-funded programming need only report those evaluations in the Evaluation Registry.

The responsibility for conducting evaluations rests primarily with bureaus and independent offices. Each bureau must perform the following functions:

- Appoint a Bureau Evaluation Coordinator.
- Ensure all its programs allocate sufficient funds to evaluate projects/programs or management processes.
- Ensure the integrity and independence of evaluations.
- Arrange for evaluation training for the staff and contractors engaged in evaluation activities.
- Work with partnering organizations to improve their expertise and capacity to evaluate interventions.
- Facilitate the utilization of evaluation findings and recommendations.
- For foreign assistance-funded evaluations, update the [Evaluation Registry in Facts Info NextGen](#).
- For diplomatic engagement-funded evaluations, update information on planned evaluations and upload documents for ongoing and completed evaluations to the [Evaluation Management System](#).
- Hold meetings and lectures for its staff to promote evaluation.
- Keep in touch with the evaluation community within ([State's Evaluation Community of Practice](#)) and outside the Department to improve its expertise in evaluation.

Budgeting for Program Design, Monitoring, and Evaluation Activities

For both diplomatic engagement and foreign assistance-funded evaluations, design, monitoring, and evaluation activities are allowable program costs that should be considered when submitting their Bureau Resource Requests. Although international standards for monitoring and evaluation range from 3 to 5 percent of the total cost of a program or project, this figure is higher than most organizations can achieve and is likely closer to between 1 and 2

percent. In the case of programs, evaluation costs should be planned and accounted for as part of the overall program budget; bureaus and independent offices must allocate sufficient resources from the program budget for evaluations when planning programs and activities. All bureaus and independent offices should also budget for conducting evaluations of their management processes and services.

Transfers of Funds

When a transfer takes place from Department of State to another agency or institution, the State bureau or independent office must obtain from the receiving institution sufficient monitoring data associated with the funds to determine if adequate progress and results are being achieved, and any evaluation findings related to the outcomes achieved with the funds. The overarching goal of these requirements is to ensure funds are being used as intended, and for bureaus or independent offices to have adequate information to understand the extent of progress toward the result(s) achieved with the transfer of funds. State bureaus and independent offices should use the information provided to learn and adapt as necessary, within the context of their broader strategic objectives, programmatic outcomes, and theories of change.

For diplomatic engagement funds, these arrangements are typically processed as reimbursements. Depending on the statutory authority under which the reimbursement occurs, the State bureau or independent office may be responsible for ensuring desired results were achieved. Consult your BP budget analyst to determine which agency is responsible for program accountability.

Transfers of foreign assistance between State and other agencies typically take place using one of two types of agreements. Per the Foreign Assistance Act:

- Section 632(a): this authority covers transfers of funds to another agency where the recipient agency takes on responsibilities for program accountability.
- Section 632(b): this authority involves interagency agreements when one agency is "buying" services from another agency, and the buying agency retains responsibilities for program accountability.

For foreign assistance, this guidance mainly applies to 632(b) transfers or other instances where State is retaining responsibility for oversight and accountability.

Before signing an interagency agreement, the transferring bureau or independent office and the recipient institution should discuss the information required to be submitted to the transferring bureau or independent office. The information to be submitted to State should be

explicitly stated in the interagency transfer agreement. In determining what information should be submitted, the transferring bureau or independent office and recipient institution should discuss what policies and practices the receiving institution already has in place to meet these aspects of 18 FAM 300 and assess what monitoring or evaluation data may already be available or is expected to be produced under the receiving institution's current policies or practices. To the extent practicable, the transferring bureau or independent office should leverage existing processes or data sources to reduce reporting burden (e.g. the use of standard foreign assistance indicators, if applicable to the transfer scope).

Bureaus and independent offices should rely heavily on the program logic models, theories of change, and performance management plan documents prepared under 18 FAM 300 when determining which information to request from the receiving agency, and consider the following questions:

- What decisions will the information provided inform, and how will it be used?
- Which data would tell us what was ultimately delivered with and/or achieved as a result of the transferred funds so we can assess whether they align with expectations?
- Which data would convey what progress is being made, and would be useful to inform/adjust ongoing operations?
- Which data could help us understand if we're doing what's articulated in broader bureau or independent office program logic models, and if the activities/deliverables are truly aligned to that logic?
- Which data could help the bureau or independent office better understand, test, or improve the viability of its theories of change or program logic models?
- Which data might be useful to inform planned evaluations, or help point out areas where evaluation might be necessary?
- How often will data need to be reported in order to be useful, and does the implementing mechanism being used support this frequency?

Bureaus and independent offices should also consider the following broader questions as they develop interagency agreements:

- Adaptation and Management: At what point would data showing less progress than expected, or more progress than expected, trigger a management decision to be made by State and the receiving agency about how to adjust?
- Evaluation: Who will decide whether to evaluate, what to ask, and when? Will the State bureau be engaged in planning for any evaluation, or will those decisions be up to the receiving agency?