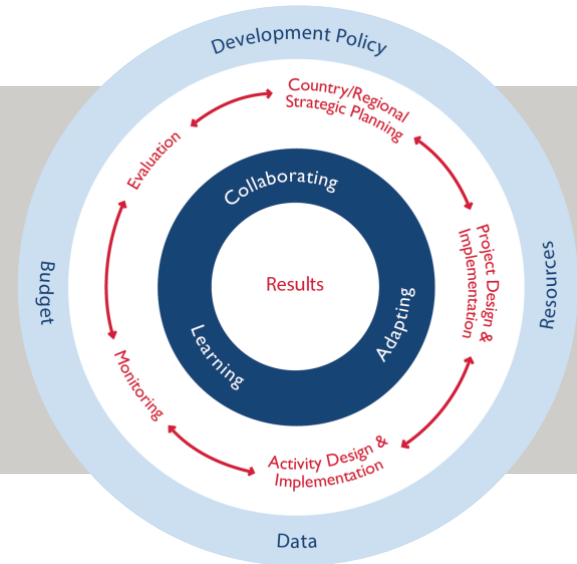# PROGRAM CYCLE

Discussion Note: Making Evidence Accessible through Evaluation Synthesis Platforms

This Discussion Note is a starting point for USAID staff interested in strengthening links between evidence from evaluations and other USAID-funded research and Agency investments in country strategies, projects, and activities.

Evaluation syntheses are analytic reviews of evaluations that aggregate findings from evaluation reports and/or examine the quality of the evaluations reviewed.

This Note describes how evaluation syntheses have been conducted to aggregate evidence in support of learning agendas and forward-looking assistance programing. It also illustrates how assessing evaluation quality and compliance with Agency policy can strengthen the credibility of the evidence USAID generates on what works and merits further investment or scaling up.

Discussion Notes explore principles or methods related to the Program Cycle and are intended to prompt inquiry. This Note was developed by the Bureau for Policy, Planning and Learning (PPL).

Introduction

Evaluation syntheses aim to consolidate what is being learned from evaluations and make that evidence more readily available. These studies can help busy decision-makers–who may not have time to locate and review all relevant evaluations and research reports–use existing evidence to better address pressing development challenges, consistent with USAID's commitment to supporting countries.

Section I of this Discussion Note describes the types of evaluation syntheses. Section II examines how syntheses can enhance evidence-based decision-making within USAID's Program Cycle. Section III discusses stages in the evaluation synthesis process, including design, implementation, dissemination, and utilization tracking. Section IV provides links to additional resources for those interested in pursuing one of the main types of evaluation syntheses or for understanding how to identify evidence gaps that learning agendas can help fill.

Section 1: Types of Evaluation Synthesis

This Discussion Note focuses on syntheses based on USAID evaluations. The term "syntheses" can also refer to studies based on other evidence sources. Twelve of the 59 verified synthesis reports in USAID's Development Experience Clearinghouse (DEC) that were published between 2012-2018 were based solely on USAID evaluations.

Figure 1 identifies two main types of evaluation syntheses. **Meta-evaluations** are evaluations of a set of evaluations. **Meta-analyses** synthesize findings from evaluations and/or other types of studies. Boxes in blue are syntheses based on evaluations, while boxes shown in grey are based on other types of evidence, or a mix of evaluations and other documents.
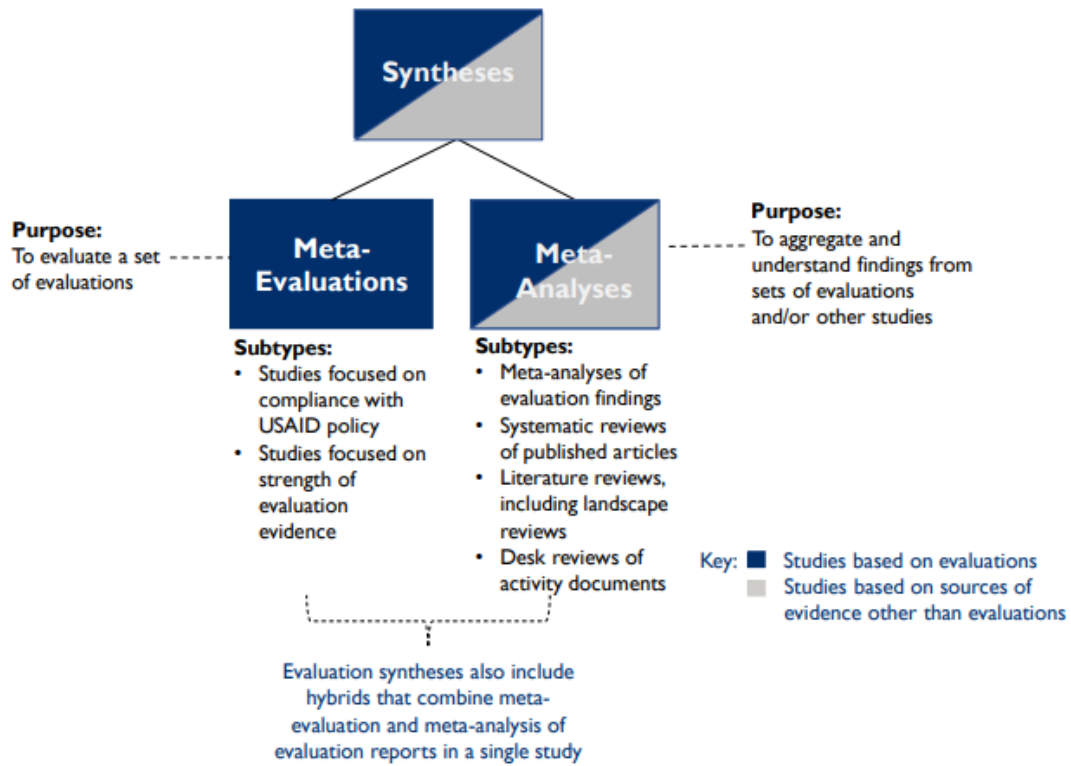
Syntheses

Meta-Evaluations

Purpose:
To evaluate a set of evaluations

Subtypes:
• Studies focused on compliance with USAID policy
• Studies focused on strength of evaluation evidence

Meta-Analyses

Purpose:
To aggregate and understand findings from sets of evaluations and/or other studies

Subtypes:
• Meta-analyses of evaluation findings
• Systematic reviews of published articles
• Literature reviews, including landscape reviews
• Desk reviews of activity documents

Key: ■ Studies based on evaluations
□ Studies based on sources of evidence other than evaluations

Evaluation syntheses also include hybrids that combine meta-evaluation and meta-analysis of evaluation reports in a single study

Figure 1 Evaluation Synthesis Types and Sub-Types

Section II: Evaluation Evidence and USAID's Program Cycle

USAID's Automated Directives System (ADS) 201 defines evaluation as, "the systematic collection and analysis of data and information about the characteristics and outcomes of one or more organizations, policies, programs, strategies, projects, and/or activities conducted as a basis for judgments to understand and improve effectiveness and efficiency, timed to inform decisions about current and future programming." This definition aligns evaluation with key Agency Program Cycle principles, one of which states that "USAID's decisions about where and how to invest foreign assistance resources should be based on analysis and conclusions supported by evidence [from] formal assessments, evaluations, and studies conducted by USAID or other development actors" (ADS 201.3.1.2)

Furthermore, ADS 201.3.2 advises on gathering and presenting evaluation evidence to support proposed strategies, projects, and activities. This means not just using evidence from evaluations conducted by an operating unit preparing a Country Development Cooperation Strategy (CDCS), Project Development Document (PDD), or Activity Approval Memorandum (AAM) but also reaching beyond a Mission's local evidence base for relevant evidence from other USAID operating units, partner governments, civil society, the private sector, and other donors. ADS 201 provides task-specific guidance for applying this principle across several Program Cycle stages.

For example:

1. In the CDCS process there is an opportunity for staff from OUs to review lessons from data from monitoring and evaluation and other sources of evidence; explore different hypotheses about how change is expected to unfold." (ADS 201.3.2)
2. Project and activity designs are to be grounded in evidence from sources including evaluations, analyses, monitoring data, etc. (ADS 201.3.4.5)
3. The importance of robust evidence for future programming decisions underlies requirements for monitoring and evaluation planning in the ADS.
4. Technical offices in USAID bureaus have a special responsibility to disseminate knowledge on recent advances and innovations in their respective technical fields, inclusive of evaluation findings, to help the Agency make evidence-based and strategic choices.

These guidelines align closely with the goal around evidence-based decision making in the current joint Department of State-USAID Strategic Plan (2018-2022).

While Agency policy encourages reaching beyond individual operating units for evidence, a 2016 report on Evaluation Utilization at USAID found that Missions often only review evaluations they commissioned and do not routinely access and build on evaluation evidence from other countries and regions where USAID is addressing similar development challenges. High-quality syntheses that are strategically disseminated can help fill that gap. Evaluation findings are also potentially relevant for updating and formulating new Agency policies. However, the 2016 utilization study found that USAID does not consistently use evaluation evidence when formulating policies.

Box 1: USAID Experience with Evaluation Syntheses

USAID's DEC archives a variety of reports and studies conducted since the Agency's creation in 1961. The DEC includes 110 documents published between 2012 and 2018 that contain one of the Box 2 terms in their titles. Fifty-nine of these 110 documents were validated as synthesis reports that conform to Box 2 definitions as shown in Figure 1. Of these 59 documents, 12 (20 percent) drew their conclusions exclusively from evaluation reports. A Technical Report available here presents findings from a 2018 review of syntheses available on the DEC and explains how a sample of 13 USAID syntheses were planned, implemented, disseminated, and utilized.

Performance Goal 4.1.1

By 2022, increase the use of evidence to inform budget, program planning and design, and management decisions.

- Department of State-USAID Joint Strategic Plan FY 2018-2022

Box 2: Definitions of Key Terms Referenced in this Discussion Note

Synthesis: Process of analytically summarizing information from multiple sources to convey their combined significance.

Meta-Evaluation: Analytic examination of a set of evaluations to assess aspects of their quality or merit. Some meta-evaluations focus on the *compliance* of a set of evaluations with the sponsoring organization's evaluation requirements and guidance. Other meta-evaluations have been commissioned to assess the *strength of the evidence* that a set of evaluations provide.

Meta-Analysis: Analytic examination of secondary source documents to identify patterns among them (convergence/divergence) also a statistical approach to combining results of multiple studies to increase their statistical power. Meta-analyses may examine evaluations as well as other types of research.

Sub-types of meta-analyses include:

- *Meta-Analysis of Findings from Multiple Evaluations*: May be topical, focusing on patterns of findings from evaluations in a specific sector. May also look across evaluations of activities in multiple sectors within a specific geographic area to understand patterns of success and failure and challenges affecting a range of interventions. (This term is sometimes used more narrowly when discussing the statistical analysis aspect of a Systematic Review, defined below).
- *Multi-Site Evaluation:* Examines a single evaluation for patterns in findings across case studies, in various locations where an intervention was delivered, or under various arms of a research trial.
- *Systematic Review*: Summarizes findings of multiple rigorous studies on a specific question or hypothesis. May examine the effect of multiple interventions focused on the same result, or examine replication studies that focus on differences in the effects of an intervention in a new environment or to additional target group types. Most systematic reviews comprehensively assemble peer-reviewed journal articles in the field of study, with screening criteria for research design quality and evidence strength. Published evaluation findings may appear in systematic reviews of interventions for which evaluations were undertaken.
- *Literature Review:* Draws on published works, but may include a wider range of documents for a topic and include documents produced by organizations working in an area as well as journal articles and other materials. It sometimes precedes work on a strategy or project design in a new area. For impact evaluations, a literature review may be a preliminary step to help establish what is already known about interventions focused on a problem or opportunity. Literature reviews are sometimes undertaken with publication as their specific goal.

Other Types of Synthesis: Other types of synthesis reports can also help build the evidence base on which decision-makers rely. Such studies do not necessarily review evaluations or published articles, although they may do so. Examples include:

- *Landscape Review:* Describes the kinds of activities being undertaken to address a problem or opportunity. It may include a description of types of activities, what agencies are supporting them, and in which regions.
- *Desk Study:* Typically includes a review of existing documentation to inform future actions. Desk studies often synthesize existing information concerning a strategy, project, or activity. They may include documents about ongoing or previous activities, the environment in which they were undertaken, and the political, economic, and other characteristics of those situations, or target groups within them.
- *Compendium of Evaluation Abstracts:* Can help quickly identify the most relevant evaluations within a particular area by distilling key aspects of evaluation reports to their essential points.

Section III: Process for Conducting an Evaluation and USAID's Program Cycle

This section provides an overview of the evaluation synthesis process. The resource sheets in Section IV provide further guidance for those interested in initiating a specific evaluation synthesis sub-type.

STAGE 1: CLARIFY THE PURPOSE, AUDIENCE, AND USES OF THE EVALUATION SYNTHESIS

Syntheses reports generally cite one of three broad purposes:

1. Increase the awareness of Agency leadership and other key stakeholders of USAID evaluation results and their practical significance.
2. Base policies and programming decisions at the Mission, bureau, or Agency levels on evidence synthesized from USAID evaluations worldwide, as well as other sources.
3. Strengthen USAID's evaluation policies and align its evaluation practice across sectors and regions.

Achieving these purposes depends on whether evaluation syntheses are *utilized*, not simply produced. Ideally, evaluation syntheses will lead to improved project and activity designs and implementation that yield better development results. The use of evaluation syntheses, in turn, depends on the strength and relevance of the evidence syntheses and how effectively that evidence is presented.

To ensure that Mission decision-makers and country partners are aware of the findings from syntheses, the following scenarios illustrate the range of Program Cycle processes for which accessing evaluation evidence could potentially enhance the analytic rigor of decisions USAID staff make.

- **Scenario 1**: A policy formulation team, comprised of staff from relevant technical support bureaus and Missions, is mobilized to update a USAID policy on assistance in a sector or subsector.
- **Scenario 2**: Mission staff have formed an internal team to prepare the Mission's second CDCS. The team must ensure the CDCS is based on existing evidence and may call on relevant technical bureaus for support.
- **Scenario 3**: A Mission has mobilized a project and/or activity design team that includes Mission staff with technical expertise and, if appropriate, USAID/Washington staff with sector expertise and experience integrating local systems thinking into project and/or activity design.
- **Scenario 4**: A Mission Project Manager has formed a team to develop a statement of work (SOW) for an activity. The activity will address a difficult and chronic problem and requires a proven, evidence-based approach to meet expectations for success.

While teams for each of these scenarios share a need for evaluation evidence, the scenarios do not indicate which type of evaluation synthesis would best support their efforts. USAID guidance does, however, suggest that Missions should consider evidence (e.g., evaluations from other sources such as other Missions, governments, civil society, the private sector, other donors) in addition to findings from their own assessments and evaluations, as relevant.

To gauge the interest of potential audiences such as USAID staff and other stakeholders in the kind of information an evaluation synthesis might provide, those considering commissioning a synthesis may find it useful to conduct a stakeholder analysis, even if on a small scale. Interacting with the potential audience for an evaluation synthesis can help clarify which questions or issues are of greatest interest for a synthesis to address.

A stakeholder analysis can also ascertain audience preferences about the structure and format of eventual synthesis reports. Without stakeholder input, commissioners of evaluation syntheses may risk underutilization of their investments. When planning a synthesis stakeholder analysis, consider how evidence-based decision making needs can also be met.

USAID's guidance for conducting an evaluability assessment prior to initiating an evaluation can inform this type of pre-synthesis analysis. Table 1, derived from the stakeholder assessment element of that guidance, illustrates how this might be incorporated into the planning stage of an evaluation synthesis.

Table 1: Sample Table for Gathering Input on Evaluation Evidence
Demand and Evaluation Synthesis Product Option Preferences

Evaluation Synthesis Stakeholder	Makes Policy	Makes Programming Decisions	Reacts to Findings	Interest Only

USAID staff, partner country stakeholders, and implementing partners will likely be able to describe the task for which evidence is needed to support decision-making and the kinds of evidence they need. However, their familiarity with specific synthesis options may not be sufficient to elicit answers about potential product types. By providing examples of relevant synthesis types, the potential audiences may better understand the main options available.

STAGE 2: SELECT A *FIT TO PURPOSE* APPROACH

Fit to purpose, a term borrowed from consumer protection law, suggests the challenges faced by those who commission evaluation syntheses. Not only must the task and evidence needs of potential users be understood, but also how diverse types of syntheses match up with those needs.

Table 2 illustrates options that might be suitable for meeting the needs of stakeholders at different points in the Program Cycle. In addition, for USAID staff considering undertaking or commissioning an evaluation synthesis, Section IV provides brief resource sheets with links to additional resources.

Table 2: Examples of *Fit for Purpose* Approach Purpose, Audience, Information Needs, and Types of Synthesis Options

Purpose to Which a Synthesis Might Contribute	Primary Audience	Specific Types of Information Needed	Synthesis Options
Prepare or update Agency-wide policy	• Policy working group members • Policymakers	• What are the priority issues in partner countries, regionally, and globally? • Where is progress being made, how, and by whom? • What is the U.S. comparative advantage by sector or region?	**Meta-analysis**, possibly: • A landscape review of relevant existing activities, including evaluations • A performance meta-analysis of relevant USAID projects/activities • A literature review for the sector or region
Update of strategic priorities for a new CDCS	• CDCS working group • Mission management	• As above, including what the country itself and other donors are doing	**Meta-analysis**, possibly: • A landscape review of relevant existing activities, including evaluations • A performance analysis of projects/activities
Present evidence on the theory of change included in a PDD	• Project design team • Mission management	• Where has this hypothesis been applied and to what effect?	**Meta-analysis**, possibly: • A literature review of published materials • A systematic review • A desk review of USAID documents, including evaluations
Prepare a Mission or USAID/Washington learning agenda or a Mission performance management plan (PMP)	• Mission or office monitoring, evaluation, and learning team • Mission or bureau management	• What is not known that could make USAID's assistance more effective? • What has been accomplished? • What activities and assumptions have failed or lagged, and why?	**Meta-analysis or a combined meta-evaluation and meta-analysis,** possibly covering: • An analysis of gaps in knowledge through a literature review or survey of practitioners • A performance analysis of projects/activities • A meta-analysis of relevant USAID evaluations • Strength of evidence if evaluations are examined
Monitor and improve evaluation quality in a Mission or bureau	• Mission or office monitoring, evaluation, and learning team • Mission or bureau management	• How well do evaluations comply with Agency evaluation policy? • How adequate is the strength of the evidence evaluation provided?	**Meta-evaluation**, possibly: • At the Mission level • At the bureau level

STAGE 3: MOBILIZE AND MANAGE RESOURCES FOR THE SYNTHESIS

During this stage, an office that plans to undertake a synthesis would decide whether to do so using Agency staff or by contracting with an external team. In either case, a SOW can help guide the team. For externally contracted teams, the office commissioning the synthesis will need to estimate the cost of the study and obtain approval for the necessary resources. Once initiated, the USAID manager for the synthesis will need to oversee the work and prepare a dissemination plan that can be carried out once the synthesis products are ready. Given the range of evaluation synthesis types, this section describes generic tasks for undertaking a synthesis. Section IV provides resource sheets for synthesis sub-types.

The duration of an evaluation synthesis varies considerably depending on sub-type. Large-scale meta-evaluations that do not include a meta-analysis element may take up to 12 months, while systematic reviews that involve comprehensive literature searches and multiple screening steps sometimes take even longer. Meta-analyses without a meta-evaluation element typically require around 6 to 12 months to complete.

GENERIC SYNTHESIS STUDY TASKS

This section describes typical steps for conducting an evaluation synthesis. While external teams are often engaged to carry out many of these tasks, USAID staff may be involved in these tasks as active participants, managers, or in hybrid roles. Integrating USAID staff into synthesis tasks can help build ownership and foster utilization of the synthesis products, although other demands on staff time often make it difficult for them to fully engage in synthesis tasks.

The Gantt chart in Table 3 summarizes the main tasks involved in a synthesis. Task descriptions are provided below the table.

Table 3: Illustrative Synthesis Gantt Chart

Tasks	Pre-Start-Up	Year			
		1st Qtr	2nd Qtr	3rd Qtr	4th Qtr
Determine Audience Needs/Interests/Product Preferences	X				
Develop Approach and Protocols		X			
Identify Data Set		X			
Train Team		X			
Implement Protocols and Supervise Team			X	X	
Data Analysis				X	

Tasks	Pre-Start-Up	Year			
		1st Qtr	2nd Qtr	3rd Qtr	4th Qtr
Report Preparation/Submission					X
Dissemination and Action Plan					X
Utilization Follow-Up					Six months later

1. **Determine audience needs, interests, and product preferences**
   This task should be undertaken before funds for a synthesis are committed. Providing examples of syntheses can help elicit audience opinions about the type of synthesis product most likely to be used. Contrasting examples might include a topic-focused meta-analysis versus a compendium of evaluation abstracts on the same topic.

2. **Develop the synthesis approach and protocols**
   Data on existing syntheses indicate that the more complex the review protocol, the longer it takes to develop and pre-test. For example, a blended quality and compliance review tool developed by USAID's Office of Education to screen out low-quality evaluations before undertaking meta-analyses took longer than the creation and pre-testing of a checklist used for an Agency-wide meta-evaluation. Meta-evaluation tools, in turn, took longer to develop than sets of questions on which various meta-analyses focused.

3. **Identify data set**
   Depending on how many evaluations are to be reviewed, one or two months may suffice for studies that extract and validate evaluations from the DEC. Studies that involve searches beyond the DEC may require an additional month or so to locate all relevant evaluations. The type of comprehensive literature review required in systematic reviews – which often start with thousands of articles identified using key terms, then screened by multiple criteria to narrow the set – can take six months or longer to reach a final set of studies. The more complex the screening process, the longer this step takes.

4. **Train team**
   Approaches for orienting and training teams vary by type of study. Studies that involve rating or scoring evaluations and necessitate ensuring inter-rater reliability among team members require more time than other team training approaches. Normally in checklist rating approaches, where all items can be rated objectively by well-trained mid-level staff, initial training on items may require around two additional weeks to achieve inter-rater reliability, depending on the number of items involved and degree of judgment required to rate them.

5. **Implement protocols and supervise team**
   Full-time supervision of teams that implement synthesis protocols is well advised. This role often involves quality checking, ensuring inter-rater reliability using regular spot-checks of multiple team members, distributing evaluations to team members, and overseeing the time taken to complete protocol tasks per evaluation.

6. **Data analysis**

   The design for an evaluation synthesis should include an analysis plan. If the synthesis team plans to use a template to record findings from the evaluations or other studies they review, that template should be included in the analysis plan, which should also explain how those findings will be compared across studies. Where appropriate, analysis plans can also include samples of the kinds of tables and graphs the synthesis team plans to include in its report. Time required for data analysis is affected by the extent of qualitative analysis required to identify patterns in findings across evaluations and establish other types of relationships based on narrative materials. A well-structured analysis plan helps foster efficiency at the analysis stage of a synthesis. Simplifying the findings and conclusions through a presentation using bullet points can help set up the report writing. The presentation can also provide the commissioning unit with a preview of the synthesis' main results, and can help reach agreement on a report outline if not already determined.

7. **Report preparation**

   Key findings and conclusions should be prominently featured at the beginning of the synthesis report. An exception is a systematic review, which should follow the style guide provided by the target journal to which the review will be submitted. Reviewing examples of previous USAID syntheses of the same sub-type may be helpful, regardless of which type of synthesis report is to be prepared. Additional dissemination products can also be created in parallel to the main study report; USAID's [two-page brief](#) on its evaluation utilization study is an example.

8. **Dissemination and action plan**

   With this task, leadership on the synthesis transfers back to the commissioning unit. A dissemination plan is important for this stage. USAID's [Evaluation Toolkit](#) provides a useful template for such a plan, which ideally was prepared earlier and with the study team's input. Ideally, USAID synthesis reports will be submitted to the DEC. However, this is not sufficient as dissemination should be active and targeted to reach intended users, often through multiple venues and formats. If the synthesis includes recommendations for USAID's action, the commissioning office should review the templates and action trackers in the [post-evaluation section of the Evaluation Toolkit.](#)

9. **Utilization follow-up**

   Following up with intended users of evaluation syntheses is the only way to understand the value and limitations of the study and ways to improve future syntheses. Some commissioning offices have good knowledge of how syntheses were received and utilized. Others, however, know little about utilization. It is in USAID's interest to learn as much as it can about utilization, usually in the first six months after a synthesis product is released. While this does not have to take a lot of time, it does require attention.

STAGE 4: PURSUE ACTIVE DISSEMINATION AND TRACK UTILIZATION

Preparing and implementing a dissemination plan and tracking utilization in a systematic way is important when investing in an evaluation synthesis. Key aids from USAID's Evaluation Toolkit include [Developing an Evaluation Dissemination Plan](#) and [Post-Evaluation Action Plans.](#)

Agency leadership have highlighted the importance USAID attaches to disseminating evaluations to country partners and other non-USAID stakeholders and, by extension, related products such as evaluation syntheses.

For all audiences, USAID encourages active forms of dissemination. Past synthesis dissemination activities include:

- Briefings to program officers;
- Alerts on sector websites;
- Infographics to draw attention to a systematic review; and
- Communications alerting targeted individuals about the existence of a new systematic review, whose findings were also presented at a national education conference.

Additional options that might also be useful for disseminating synthesis findings include:

- A two-page briefing note, like those PPL/LER has posted on its [webpage;](#)
- Social media alerts on Twitter and other messaging sites;
- A blog post on USAID's Impact blog page;
- Webinars in which USAID staff overseers can participate directly; and
- A short YouTube video presentation or interview with synthesis authors.

While utilization tracking was not significant in past post-study work plans, some commissioning units were aware of important utilization results such as:

- The use of a Bureau for Resilience and Food Security [evaluation synthesis](#) in the U.S. government's latest Global Food Security Strategy; and
- Changes in a Mission's SOW template as a result of a synthesis.

Section IV: Resource Sheets

This section provides resource sheets for conducting evaluation syntheses: meta-evaluations, meta-analyses, and rigorous systematic reviews.

Resource Sheet for Conducting Meta-Analyses

The term meta-analysis is used in several ways. Empirically, a review of studies termed meta-analysis in the international development arena indicates a continuum of studies that apply a structured approach to assemble and analyze evaluations as well as other types of research to extract findings on research questions, topics of interest, cross-cutting issues, and other matters of interest to program decision-makers. At one end of this continuum are literature reviews. Studies that focus on narrative or other types of qualitative data are sometimes called qualitative meta-syntheses in published journal articles but are almost always called meta-analyses or syntheses when conducted at USAID's behest. At the other end of the meta-analysis continuum are systematic reviews, for which a separate resource sheet is provided in this Discussion Note. These studies also review existing literature, but screen entries on quality criteria before examining them. Among organizations that conduct systematic reviews, the term meta-analysis is sometimes reserved for describing the statistical procedures used in systematic reviews to review the effect size of an intervention across several similarly focused randomized trials (or USAID impact evaluations).

Guidance specific to meta-analysis across this continuum does not exist in a single volume. Instead, there are separate guides for conducting literature reviews, qualitative meta-syntheses, and systematic reviews (provided in the following resource sheet). For literature reviews, the Overseas Development Institute's How to Do a Rigorous, Evidence Focused Literature Review in International Development is useful, somewhat spanning the continuum above. Writing on qualitative meta-syntheses is relatively new, with perhaps the most generally useful summary of this synthesis type provided in a 2018 conference paper that describes the type of study covered by this term by contrasting them with systematic reviews.

The high percentage of syntheses conducted by USAID that fall under the term meta-analysis, irrespective of whether they include evaluations, indicate a strong interest in bringing together evidence and lessons on development assistance interventions based on the experience of USAID and other sources of development assistance. Dissemination of those that exist is equally important, as it is important for USAID staff to be aware of existing meta-analyses that may be pertinent to the projects and activities the Agency is planning. A recent critique suggests that may not always be the case.

Resource Sheet for Conducting a Systematic Review

Systematic reviews are analytic reviews that summarize the findings of multiple studies on a specific question or hypothesis. They are frequently used to synthesize results of multiple randomized controlled trials. In the international development arena, 3ie has conducted many systematic reviews, covering fields for which a sufficient number of rigorous evaluations and other types of unbiased research trials have been completed. USAID, DFID, and other donors have helped underwrite these studies. Domestically, the Cochrane Collaboration is a leader in conducting systematic reviews in health, while the Campbell Collaboration conducts them for education, justice, social welfare, and other fields, including studies on international development. Among the 30 systematic reviews USAID conducted between 2012 and 2018, most have been health related and several have been published in peer-reviewed journals.

Guidance on how to produce a systematic review is available from several sources, including the Cochrane Collaboration. One of the most approachable guides comes from Cornell University. The Cornell Guide walks readers through steps in the process and includes the useful comparison between traditional literature reviews and systematic reviews reproduced below. Because of their comprehensive nature, the initial steps involved in assembling and screening studies is lengthier than for other synthesis types, with two years being well within the norms for this step.

Resource Sheet for Conducting Meta-Evaluations

Meta-evaluations generally focus on aspects of evaluation quality, ranging from compliance with evaluation policies of the organizations that commission them, to widely accepted measures of evaluation strength (e.g., validity, reliability, generalization). USAID, UNICEF and several bilateral donors have undertaken meta-evaluations, using them as a tool for improving evaluation practice.

A generic meta-evaluation checklist for reviewing evaluation reports based on standards set by the American Evaluation Association provides a broad view of the coverage of this type of study. The current USAID version is available in the Agency's Evaluation Toolkit, and is recommended as a tool for reviewing draft evaluation report compliance with policy. A parallel USAID checklist for evaluating evaluation SOWs could serve as a basis for conducting a meta-evaluation of evaluation SOWs.

Some meta-evaluations focus on evidence strength rather than policy compliance. GRADE is a leader in this area, representing a collaboration of organizational efforts to rank evidence quality in clinical studies, and meta-evaluations examining elements of evidence strength are routinely embedded in systematic reviews as screening devices. With similar intent, USAID's Office of Education developed and applied an evidence strength rating tool to screen evaluations for inclusion in meta-analyses it conducted in 2018, basing this work on efforts of a broader donor collegium to identify evidence strength dimensions in education. USAID's compendium of evaluation abstracts for mobiles in education used an abbreviated evidence strength rating process by simply classifying abstracts based on the type of evaluation design employed.

Standard steps in a meta-evaluation process include:

- Developing a checklist or other instrument to be used, recognizing that objective criteria reduce the likelihood of variance among reviewer ratings or the chances for bias in ratings;
- A sampling plan for selecting a representative set of evaluations (e.g., by organization or unit, year, sector);
- Mobilization, training and supervision of reviewers, with a focus on inter-rater reliability; and
- Analysis and reporting.

Existing meta-evaluations by development assistance organizations vary as to whether and how they transform checklist approaches into scores and monitor improvements over time. The time and cost of a meta-evaluation varies by the number of evaluations reviewed, as well as by the skill level of the personnel involved. Meta-evaluations based on compliance checklists can be implemented by junior- or mid-level staff, while the evaluation skills and experience needed to implement meta-evaluations based on evidence strength can be significantly higher. Relatively large meta-evaluations that examined hundreds of evaluations have taken one-to-two years to complete.

References:
Cooksy, L., & Caracelli, V. (2009). Metaevaluation in Practice. *Journal of MultiDisciplinary Evaluation,* 6(11), 1-15.

Hedler, H., & Gibram, N. (2009). The Contribution of Metaevaluation to Program Evaluation: Proposition of a Model. *Journal of MultiDisciplinary Evaluation,* 6(12), 210-223.