



USAID
FROM THE AMERICAN PEOPLE

ENHANCING USAID IMPACT EVALUATION INVESTMENTS

Action Recommendations

Updated March 2021

USAID's 2011 evaluation policy and its 2016 update increased the Agency's attention on impact evaluations (IEs). The number of USAID-funded IE reports published each fiscal year (FY) rose from 6 in 2012 to a peak of 32 in 2018. In 2019, USAID's Office of Planning, Learning, and Coordination in the Bureau for Economic Growth, Education, and Environment (E3), in collaboration with the Office of Learning, Evaluation, and Research in the Bureau for Policy, Planning, and Learning (PPL/LER), commissioned a review of its completed IE reports to assess their consistency with Agency policy and guidance and with professional norms of IE quality.¹ The review identified these reports' strengths and weaknesses and provided recommendations to improve the quality of IE reports at USAID. The review's recommendations complement the lessons and tools identified in a prior discussion paper, *Impact Evaluation: Critical Challenges/Promising Solutions*.² Shortly after the IE quality review was completed, USAID released an update to its Automated Directives System (ADS) 201 guidance³ that aligned with some of the review's recommendations.

This briefing note summarizes key results from the IE quality review and expands on the review's recommendations to reflect the ADS 201 updates and operationalize them for USAID stakeholders responsible for commissioning and managing IEs. It also provides an annex with relevant resources discussed in the recommendations.

Impact Evaluation Quality Review

The review assessed the quality of all 133 USAID-funded IE reports published between FYs 2012-2019. It first assessed whether the report met USAID's definition that IEs "require a credible and rigorously defined counterfactual." For reports that met this initial criterion, the review then assessed whether they addressed quality-relevant elements across six domains: sample size considerations, conceptual framing, treatment characteristics and outcome definitions and measurement, data collection and analysis, common threats to validity, and reporting of findings. The elements within each domain were separated into a first tier (basic quality elements) and a second tier (additional quality elements) to avoid arbitrarily weighting the items differently based on perceived importance.

The review found that 54 percent of IE reports met USAID's IE definition, 28 percent failed to provide statistical justification for the validity of the comparison group, and 18 percent did not have a comparison group. Based on the first-tier score, 17 percent of the IE reports were of high quality, 30 percent were of acceptable quality, and 7 percent were of lower quality. Second-tier scores showed that only 3 percent of IE reports were of high quality, 14 percent were of acceptable quality, and 28 percent were of lower quality.

While the quality of IE reports improved over time, particularly on basic quality elements, critical gaps need to be addressed to improve their quality. First, two to four IE reports each year still do not meet USAID's IE requirement for a comparison group. Second, IE reports – in particular those with non-equivalent groups design – do not consistently provide statistical justification for the comparison group's validity, thus compromising attribution evidence. Third, the IE reports typically did not include quality elements that were not expected in performance evaluations (and thus not explicit in USAID's evaluation reporting guidance).

¹ The E3 Analytics and Evaluation Project conducted the review, which can be found at: https://pdf.usaid.gov/pdf_docs/PA00X78R.pdf.

² The discussion paper examined technical and management challenges USAID staff face when undertaking an IE and presented tools and solutions to address these challenges. See Hageboeck, Molly, Jacob Patterson-Stein, and Irene Velez. "Impact Evaluation: Critical Challenges/Promising Solutions." Management Systems International, August 2019, https://pdf.usaid.gov/pdf_docs/PA00WFBB.pdf.

³ ADS 201 was updated on October 28, 2020. See <https://www.usaid.gov/ads/policy/200/201>.

However, not all quality elements that are explicit in USAID guidance were prevalent in IE reports. These findings imply a gap not only in IE-specific guidance but also in compliance with general evaluation reporting requirements. Fourth, most IE reports merely noted impact estimates and statistical significance without explaining *why* there was an impact or lack thereof. Doing so requires articulating a theory of change and defining outcomes along the causal pathways, incorporating qualitative methods, conducting implementation fidelity monitoring, addressing common threats to validity, and explaining null effects. Fifth, IE findings need to be accessible and discuss the practical significance of impact estimates to allow users to interpret whether the effect magnitude is large enough to be of practical value. Such information enables comparison to reported effects from other IEs and can help USAID make decisions on resource allocation. Sixth, 40 percent of IEs did not report ethical considerations, which is standard practice for research involving human subjects. Finally, few IEs (15 percent) conducted cost-effectiveness analysis to link an intervention's effectiveness to its costs.

Action Recommendations

The recommendations below have been updated since the IE quality review report to account for recent revisions USAID made to ADS 201 and have been operationalized for three types of USAID IE stakeholders: PPL/LER, IE managers, and technical officers. These recommendations, along with the tools and lessons learned from the prior IE discussion paper,⁴ propose solutions to help Agency staff successfully manage IEs and enhance USAID IE investments so they generate high-quality, credible evidence for use in decision making. The relevant tools and resources are provided as an annex to this briefing note.

Recommendations for USAID/PPL/LER

- 1. Reinforce that IEs must include a comparison group.** The recent ADS 201 revisions expand on USAID's IE definition by explicitly stating that IEs *must* use an experimental or quasi-experimental design, both of which require comparison groups. However, this revision alone may not be sufficient to ensure that future IEs comply with this requirement. PPL/LER should socialize this requirement to technical and program officers, targeting operating units that have upcoming IEs planned, and encourage them to use this criterion to decide whether an IE is most appropriate to support their evidence needs. This can be achieved in part by continuing previous support efforts like the [Impact Evaluation Clinics](#) PPL/LER has co-hosted and updating and disseminating decision-oriented resources like the [IE decision tree](#) and the [Deciding to Evaluate Flowchart tool](#).
- 2. Provide updated detailed guidance on specific elements that should be included in IE reports.** Current USAID guidance on evaluation report structure and content ([ADS 201maa](#) and [ADS 201mah](#) as of March 2021) does not include elements specific to IEs. PPL/LER should update and disseminate detailed guidance requiring that final IE reports include:
 - Statistical justification of the comparison group's validity;
 - Explicit evaluation questions linked to the evaluation purpose;
 - A theoretical framework including a literature review, theory of change, and hypotheses to be tested;
 - Defined and operationalized outcomes in the methodology section, before presenting findings;
 - Specific power calculation parameters;
 - Detailed and complete information on common threats to validity;
 - Reporting findings that include point estimates and statistical significance as well as the comparison group mean to interpret the effect's magnitude;
 - Discussion of null effects; and
 - Actionable recommendations that advise decisionmakers on the implications of the IE's results.

⁴ Hageboeck et al. (2019), https://pdf.usaid.gov/pdf_docs/PA00WFBB.pdf.

3. **Develop a standard IE-specific report template and checklist to foster inclusion of these quality elements.** USAID’s [evaluation report template](#) and [evaluation report checklist and review template](#) are useful but, as of March 2021, are not specific to IE reports. PPL/LER should adapt the instrument from the IE quality review into an IE report checklist and share it across the Agency. While it may require additional effort by USAID evaluation managers and evaluators to review IE reports against the checklist and to document more detailed information, the cost of doing so is low compared to funding an IE that does not provide useful or credible results because of poor reporting.
4. **Provide guidance to encourage IEs to report more information that disentangles and explains effects (or the lack thereof).** Most IE reports focus on answering whether an activity/intervention is effective in achieving a specific outcome. PPL/LER should provide guidance on designing IEs to also answer *why* there is or is not an impact. IEs can be designed to disentangle effects by testing combinations of interventions or to account for or measure variance in interventions or in their delivery. In addition, measuring along the causal pathways, monitoring implementation fidelity, and using qualitative methods can help explain effects or the lack thereof.
5. **Integrate ethics considerations as an IE standard to align with USAID’s Scientific Research Policy and operationalize updated ADS guidance.** The Agency’s [Scientific Research Policy](#) requires that USAID-funded research conform to host country legal and other requirements governing research with human subjects and that a study must be reviewed and approved, or deemed exempt, by a U.S.-based institutional review board (IRB). However, ethics considerations appear explicitly only in the recent ADS 201 revisions, which state that evaluations should “safeguard the dignity, rights, safety, and privacy of participants and other stakeholders and affected entities” and that “evaluators should abide by current professional standards and legal requirements that pertain to the treatment of participants.” PPL/LER should operationalize this requirement by ensuring that IEs obtain informed consent from participants and receive U.S.-based IRB approval and host-government ethics approval. The annex of this briefing note provides [Resources on Ethics Considerations](#).
6. **Develop guidance and provide support on planning and conducting cost analyses within IEs.** The ADS 201 update states that “all impact evaluations must include a cost-analysis of the intervention or interventions being studied.” Since cost analysis is not frequently conducted within IEs, PPL/LER should provide institutional support to technical officers on how implementing partners (IPs) should report cost data and to evaluation managers on how evaluators should plan for and conduct cost-effectiveness analysis. PPL/LER should also develop guidance on how to conduct this analysis, including planning and timeline implications, required expertise, coordination between evaluators and IPs, and methods. The annex of this briefing note provides [Resources on Cost Analysis](#).

Recommendations for USAID Evaluation Managers and Units Commissioning IEs

1. **When planning and designing an IE, ensure that it includes a valid comparison.** To comply with the requirement that IEs include “a credible and rigorously defined counterfactual,” the Agency’s evaluation managers and operating units commissioning an IE should use this criterion as a first screen to decide – in both the planning and design phases – whether an IE is feasible. In the planning phase, an evaluability assessment provides an opportunity to determine if a valid comparison group can be constructed. In the design phase, a useful mechanism for addressing evaluability feasibility issues early on is to conduct an IE design workshop and scoping visit with the evaluation team and the IP soon after activity award. USAID should only fund adequately powered IEs with valid comparison groups; otherwise, the evaluation questions could be revised to conduct a performance evaluation. The annex of this briefing note provides [Resources on Evaluability Assessments](#).
2. **Commission external peer reviews to assess the quality of evaluation designs and final reports, especially when there are gaps in internal technical capacity to adequately do so.** The additional reporting requirements and evaluability assessments proposed may require technical expertise

that is not always available within an operating unit. External peer reviews of evaluation designs ensure that plans for executing the IE are free of serious framing flaws and other technical impediments to the IE's success. Peer reviews of IE reports can include use of the proposed report checklist as well as a questionnaire to guide reviewers. Several USAID IEs have conducted external peer reviews but the Agency does not have a formal process or protocol for conducting such reviews of IE designs. The annex of this briefing note provides [Resources on External Peer Reviews](#).

- 3. Integrate implementation fidelity monitoring into IE statements of work.** Implementation fidelity monitoring can help address and account for common threats to validity. It can also provide insights into why impacts did or did not occur. However, gathering this information requires resources and early coordination with activity IPs. Including implementation fidelity monitoring in IE statements of work helps ensure adequate budget, timeline, and planning to gather this information. The annex of this briefing note provides [Resources on Implementation Fidelity Monitoring](#).
- 4. Ensure evaluators integrate ethics considerations as an IE standard.** To comply with the revised ADS 201 evaluation principles and standards, USAID evaluation managers should ensure IE data collection protocols include gathering informed consent from participants and should support evaluators in obtaining ethics approval from in-country IRB entities, if applicable. USAID missions could also secure a memorandum of understanding with the in-country IRB entity to streamline the ethics approval process when multiple IEs and other research activities are planned. The annex of this briefing note provides [Resources on Ethics Considerations](#).
- 5. Request that evaluators plan for and conduct cost-effectiveness analyses in all future IEs.** ADS 201 now requires that all IEs include a cost analysis of the intervention being studied. Evaluation managers should ensure evaluators incorporate cost-effectiveness analysis into their IE design protocols, budget, timeline, and team capabilities. The annex of this briefing note provides [Resources on Cost Analysis](#).

Recommendations for USAID Technical Officers

- 1. Include language in contracts/agreements on key elements that may be affected by the planned IE during the launch and implementation of activities.** ADS 201 already requires that activity planners work with A/CORs to include language about planned IEs in activity solicitations as relevant. Technical officers should also include language in contracts about expected coordination with the evaluator to define a comparison group, wait to start activities until baseline data are collected, support collecting implementation fidelity monitoring data, and report cost data for cost-effectiveness analyses. The annex of this briefing note provides [Resources on IE Language in Contracts/Agreements](#).
- 2. Facilitate coordination and communication between IPs and evaluators throughout the IE.** Technical officers should support evaluation managers in connecting IPs and evaluators as soon as the activity is awarded. This should include initial calls to discuss the IE and activity plans, an early IE design workshop between evaluators and activity IPs, and scoping activities (if needed) to assess the feasibility of IE design options. In addition, technical officers should facilitate consistent communication between IPs and evaluators throughout implementation by sharing quarterly and annual reports and workplans, implementation updates, and any changes to activity implementation that may affect the IE. The annex of this briefing note provides [Resources on Coordination and Collaboration](#).
- 3. Provide information and support to program officers during evaluability assessments and peer reviews of evaluation designs and reports.** As activity A/CORs, technical officers are in a crucial position to provide evaluation managers with useful information regarding the feasibility of IE designs and how the IE may affect activity implementation (and vice versa). Technical officers should support evaluation managers to better understand whether the IE design will result in an adequately powered study to measure changes with a valid comparison group. This entails providing information on any challenges or threats to how the comparison group will be constructed, whether IPs can maintain fidelity to the IE design, and whether sufficient units can be included in the IE.

Annex: Resources

“Annex A: Resources for Impact Evaluation Managers” in the [Impact Evaluation: Critical Challenges/Promising Solutions](#) discussion paper provides more details on these resources and additional useful tools to successfully manage IEs.

Resources for Evaluability Assessments

Planning Phase

- PPL/LER guidance on [Conducting an Evaluability Assessment](#) (2015)
- USAID’s Learning Lab also offers the Overseas Development Institute guide, [Evaluability Assessment for Impact Evaluations: Guidance, Checklists, and Decision Support](#) (2015), as a resource for conducting this process
- [Evaluation Stakeholder Analysis](#)
- [How to Note: Developing a Project Logic Model](#)
- Gugerty, Mary Kay and Dean Karlan, “[Ten Reasons Not to Measure Impact—and What to Do Instead](#),” *Stanford Social Innovation Review* (2018)

BOX 1: EVALUABILITY ASSESSMENT FOCUS QUESTIONS

1. Would an impact evaluation be useful and used?

- Are there specific stakeholder needs that an impact evaluation would satisfy and can a study be designed to meet those needs in a timely manner?

2. Is it plausible to expect impacts?

- Do stakeholders share a clear understanding of how the program operates and are there logical links from program activities to intended impacts? Does a “theory of change” depicting these links exist?

3. Is it feasible to measure impacts?

- Is it possible to measure the intended impacts of the activity, given the resources and timeline available for the impact evaluation and for the program implementation strategy?

- Adapted from *Planning for Cost-Effective Evaluation with Evaluability Assessment* (2008)

Design Phase

POST-AWARD IE DESIGN WORKSHOP

Post-award IE workshops can be hosted by USAID missions or other operating units for IPs and evaluation teams. These workshops take place before scoping and site selection visits occur or work on the IP’s monitoring, evaluation, and learning plan begins. If two CORs/AORs are involved – one for the activity implementation and another for the IE – it is important for both to participate.

The post-award workshop should take place as early as possible upon award of both mechanisms. A structured approach for these workshops (TABLE 1 provides a sample agenda) contains several standard elements, including:

- An orientation to IE methods for the activity IP, mission staff, and country stakeholders, if appropriate.
- A discussion of the purpose of the assistance activity; USAID’s hypotheses about the intended activity outcomes; the IP’s strategy, plans, and schedule for delivering the activity; and questions about activity effects that the IE might be able to answer.

- The evaluation team's initial ideas about how to design the IE, particularly around constructing a counterfactual that could support attributing observed changes in outcomes of interest to the activity's interventions. If a theory of change diagram does not exist for the activity, this workshop is a crucial time to develop and reach agreement on one.
- A discussion about the expected coordination during activity implementation, the roles and responsibilities of all evaluation stakeholders, and a review of the current timeline and next steps.

By the end of a post-award IE workshop, or through follow-up discussions, it is helpful for stakeholders to agree on and document their roles and responsibilities, as well as communication protocols, over the course of the evaluation. The importance of clarity concerning roles and responsibilities increases as the number of actors rises. Impact evaluations of mission-funded activities that are undertaken by or with the help of a USAID/Washington bureau are inherently complex, as such arrangements may involve two USAID activity managers (one for the activity and another for the IE), as well as two external entities that each may report to a different activity manager.

TABLE 1: ILLUSTRATIVE OUTLINE FOR A POST-AWARD IMPACT EVALUATION WORKSHOP

Time	Day 1	Day 2
9:00 – 10:30	<ul style="list-style-type: none"> Welcome Introductions Workshop Objectives Agenda Review <p>Discussion: Why evaluate? What types of questions can different evaluations answer?</p>	<p>Evidence Review</p> <ul style="list-style-type: none"> What has previous research in the sector shown? What are some ongoing evaluations in the sector? What isn't known – evidence gaps? <p>Group Session:</p> <ul style="list-style-type: none"> ➤ Refine Evaluation Purpose and write hypothesis as Evaluation Questions
	Break	Break
11:00 – 12:30	<p>Impact Evaluation Introduction</p> <ul style="list-style-type: none"> Impact Evaluation Methods Impact Evaluation in Practice 	<p>Theory of Change</p> <ul style="list-style-type: none"> Brief presentation guiding Theory of Change <p>Group Session:</p> <ul style="list-style-type: none"> ➤ Develop Theory of Change
	Lunch	Lunch
1:30 – 3:00	<p>Project Overview Presentation</p> <ul style="list-style-type: none"> Project Description and Implementation Plans 	<p>Outcomes Measures Discussion</p> <ul style="list-style-type: none"> Identifying the dependent variables for an impact evaluation and sources of data <p>Group Session:</p> <ul style="list-style-type: none"> ➤ Revise outcome measures of interest and data sources
	Break	Break
3:30 – 5:00	<p>Group Session:</p> <ul style="list-style-type: none"> ➤ Evaluation Purpose: What are the goals for this evaluation? ➤ Specific hypotheses to be tested: (What do we want to learn from this project?) <p>Wrap-Up</p>	<p>Group Presentation Overview Discussion</p> <p>Group Session:</p> <ul style="list-style-type: none"> ➤ Continue working on linking outcomes to TOC to Evaluation Questions ➤ Start discussion on target population, eligible beneficiaries, and plan for roll-out of activities <p>Wrap-up</p>
	Day Ends	Day Ends

Time	Day 3	Day 4
9:00 – 10:30	<p>Compared to What?</p> <ul style="list-style-type: none"> • Why a counterfactual matter • How to randomize (unit of treatment assignment) • Other options: How equal are the target (treatment) groups and the counterfactual (non-treatment) groups? <p>Group Session: Is randomization feasible?</p> <ul style="list-style-type: none"> ➤ Evaluation Design Decision Tree 	<p>** Group Presentation of Evaluation Design **</p>
	Break	
11:00 – 12:30	<p>Threats to Impact Evaluation</p> <ul style="list-style-type: none"> • Sample size considerations • Attrition, low take-up • Contamination and spillovers <p>Group Session:</p> <ul style="list-style-type: none"> ➤ Identify potential threats to the impact evaluation ➤ How can these threats be minimized or prevented? 	<p>Next Steps for the Impact Evaluation Design – Discussion:</p> <ul style="list-style-type: none"> • Scoping activities: Uncertainties in the feasibility of impact evaluation design • Coordination between impact evaluation and implementing partner team • Baseline before the intervention starts – treatment and comparison groups
	Lunch	
1:30 – 3:00	<p>Group Session:</p> <ul style="list-style-type: none"> ➤ Evaluation Design: Is randomization feasible? Why or why not? Key factors to enable randomization? Other options? ➤ Minimizing threats to the validity of the evaluation 	<p>Next Steps for the Impact Evaluation Implementation Stage – Discussion:</p> <ul style="list-style-type: none"> • Parallel Timelines: Creating a detailed schedule • Coordination between impact evaluation and implementing partner team
	Break	
3:30 – 5:00	<p>Group Session:</p> <ul style="list-style-type: none"> ➤ Finalizing Questions and Evaluation Design Method ➤ Beyond an impact evaluation: Other questions of interest and how these can be answered (performance evaluation component)? <p>Wrap-up</p>	<p>Dissemination Plan:</p> <ul style="list-style-type: none"> • Sharing evaluation findings: What? Who? When? How? • Timeline of important policy milestones <p>Wrap-up and Next Steps</p>
	Day Ends	Workshop Ends

SCOPING VISIT TO VALIDATE IE FEASIBILITY

Another important step in the start-up phase of an IE involves the finalization of decisions about where the intervention to be studied will be implemented and how a valid counterfactual will be constructed. Both the activity IP and the evaluation team may have ideas about treatment sites and populations, or how they should be selected, by the time they meet. Site selection procedures, as well as the feasibility of other aspects of IP and evaluation partner plans, thus are often a focus of initial workshops and consultations.

It is beneficial for an activity IP and IE teams to join forces for their initial scoping trips. Working in parallel in the field helps produce a shared understanding of the context in which each partner will carry out their separate but related tasks. Scoping visits carried out collaboratively by evaluation teams and IPs have proven effective in highlighting aspects of activity and evaluation planning where ideas converged, as well as pointing out where disagreement existed, thus providing a basis for discussion about areas for compromise.

As shown in the illustrative list in TABLE 2, a scoping visit, either alone or with an IP, offers evaluators an important opportunity to “ground truth” and refine initial ideas about the optimal IE design and methods.

TABLE 2: ILLUSTRATIVE IMPACT EVALUATION SCOPING VISIT TASKS

<ul style="list-style-type: none">• Determine the feasibility of a counterfactual• Assess impact evaluation design options• Select an assignment method• Estimate likely attrition	<ul style="list-style-type: none">• Determine sample size and power• Evaluate local survey research capacity• Estimate logistical challenges• Initiate institutional review board process, if needed
---	---

The number of design features to check during a scoping visit will vary from study to study. When both the IP and evaluation teams participate, it is best for both teams to approach the scoping visit with an open mind and discuss their views on various issues that arise in a collaborative manner, seeking the best overall solutions.

- In one activity, the evaluator may require data from individual units (e.g., families, farms, school children) but due to the nature of the intervention perhaps the treatment assignment unit is at a cluster-level (e.g., villages, classrooms), which would drive up an evaluation’s required sample size and may stretch the resources of the implementing partner.
- In another activity, the presence of conflict in the region and the IP’s draft work plan may block the evaluator’s hope of randomly assigning villages to treatment and control groups, forcing the evaluator to choose among quasi-experimental options. Other factors, such as conflict-induced mobility, may drive up likely attrition estimates, again raising the sample size estimate, and logistical challenges such as seasonal issues may stretch the amount of time required to complete data collection rounds. In addition, an evaluation team may learn that institutional review board clearance is required prior to initiating field research, and that it will need to hire a local expeditor who can represent the team at various stages of this process.

Each situation will likely be unique and by the end of the scoping visit, the evaluator may have a plan for executing the IE that is solid but differs greatly from what was envisioned prior to the scoping trip.

Resources for External Peer Reviews

External peer reviews of evaluation design proposals and draft evaluation reports can help ensure that IEs comply with established evaluation standards and principles and can provide technical recommendations that strengthen the IE’s validity so the results contribute to the evidence base in the respective sectors.

The E3 Analytics and Evaluation Project developed an assessment tool (provided below) for peer reviewers to analyze whether the evaluation design was feasible, methodologically sound, and informed by existing evidence and to recommend improvements or alternative options for consideration. For each evaluation design

proposal, the Project selected two peer reviewers based on IE expertise and/or sectoral expertise. The reviewers prepared a written critique and recommendations using the assessment tool. The Project also developed an assessment tool for peer reviews of draft evaluation reports (also provided below), following the same process, but in addition included criteria on fidelity to the evaluation design, data analysis, and the presentation and interpretation of results.

Impact Evaluation Design Peer Review Template

Please answer each of the questions below to address where the Evaluation Design Proposal is successful in meeting these criteria, but more importantly where the Proposal needs to be improved or expanded on, as well as, any additional options or alternatives that should have been considered.

Each question below should be answered concisely but thoroughly. Please elaborate on key points using your own expertise or past studies that have provided insights or challenges into different components of the evaluation design.

Peer Reviewer Name: _____

Date: ___ / ___ / ___

Motivation for the Impact Evaluation
1. Is the impact evaluation designed to address a specific policy problem so that results can be used to inform future decision-making and/or programming?
2. Does the Evaluation Design Proposal show where the evaluation is situated in the existing literature and how it will contribute to the current state of evidence?
Evaluation Questions
3. Are the main evaluation questions stated as testable hypotheses about cause-and-effect relationships between the program and the anticipated outcomes?
4. Are all of the evaluation questions consistent with the theory of change?
Theory of Change
5. Does the theory of change clearly outline the causal pathways from the interventions to outputs to intermediate and final outcomes?
Key Outcome Variables
6. Is each outcome variable adequately specified (i.e. clearly defined to gather reliable data that measures the construct they were designed to measure)?
7. Are important intermediate outcomes, within the scope of the evaluation, included to help explain the underlying causal mechanisms? Are there any missing outcomes that should be included?
Evaluation Design
8. If this is a randomized experiment, does the design ensure that randomization will be conducted at the appropriate level?
9. If this is not a randomized experiment, does the quasi-experimental approach proposed select the comparison group in such a way that raises concerns about selection bias? Are there alternative methods that could be considered that would reduce selection bias further?
10. Does the evaluation design raise concerns that spillovers or contamination into the control group during the course of implementation could compromise the evaluation?

11. If the intervention is one in that involves selection into treatment (for example, voluntary take-up) does the evaluation methodology account for this?
Power Calculations and Sample Size
12. Is the appropriate approach to calculate power or sample size used given the evaluation design and outcomes measures of interest? Are estimates for key parameters explicit and consistent with typically used standards?
13. Is the expected minimum detectable effect size reported? Is the basis for this estimate explicit and justified?
14. If this is a clustered design, do the power calculations account for the intra-cluster correlation (ICC)? Is the basis for this estimate explicit and justified?
15. Do the power calculations explicitly account for imperfect take-up and attrition, when appropriate?
Implementation Fidelity and Monitoring Plan
16. Is there a proposed plan to monitor implementation fidelity, compliance, and external events to maintain integrity of the evaluation design?
Data Collection Methods
17. Are the data collection methods appropriate for each of the outcome measures of interest?
18. Is the qualitative component likely to enhance the quantitative component? Does the scope of the qualitative component seem like the best approach for the study?
Data Analysis Plan
19. Is the estimation strategy (i.e. intent-to-treat or treatment-on-the-treated) and treatment effect equation (i.e. difference-in-differences, ANCOVA, post-specification) clearly laid out, including additional methods to conduct appropriate robustness checks?
20. Does the Evaluation Design Proposal include discussion of appropriate statistical corrections including control variables, standard errors, and clustering?
21. Does the analysis plan for integrating quantitative and qualitative data and findings seem adequate?
Timeline
22. Is the timeline for implementation of the project and evaluation aligned so that a) baseline data can be collected before the interventions start and b) enough time passes to expect changes in the outcome variables before endline measurements are collected?
Limitations
23. Based on this review, are there any serious risks and/or limitations that should be addressed before moving forward with the evaluation?
Publication
24. Does this evaluation design meet the standards for publication?
Additional Comments
25. Do you have additional comments on other important factors?

Impact Evaluation Draft Report Peer Review Template

The primary goal of the Impact Evaluation Report is to provide USAID with information about the impact of its programs for evidence-based decision-making. Through a Peer Review process, USAID would like to gather expert opinion about

the quality of this evaluation report. The following document contains a series of questions reflective of USAID criteria representative of a technically sound and successful Evaluation Report.

As the reviewer, you are being asked to answer the questions below to the best of your ability, honestly, and concisely but thoroughly. The feedback you provide will inform the Agency on the strengths and weaknesses of the report to ensure the findings are applied appropriately.

Please keep in mind that the evaluation report was prepared for a general USAID audience, rather than a technical or academic one, and therefore may not contain the level of detail required for a peer-reviewed publication.

Peer Reviewer Name: _____

Date: ___ / ___ / ___

Utilization for the Impact Evaluation
1. Does the impact evaluation’s design and analytical approach appropriately address the evaluation questions so that results can be used to inform future decision-making on programming? Please elaborate if yes or no, and what could have been better approaches or methods.
Evaluation Design Implementation
2. What are the key strengths and weaknesses in the evaluation design and methodology? Please elaborate your answer for each of these components: <ul style="list-style-type: none"> a. Design: b. Data Collection Methods: c. Analytical Models:
3. Does the report provide sufficient information to assess whether attrition or incomplete participation into treatment pose a threat to the evaluation’s estimation of impact? Are details on sample size across data collection rounds and their implications for statistical power clearly stated?
4. Does the report specify adequate steps taken to mitigate common threats to the integrity of the evaluation, such as non-equivalence at baseline, non-compliance, and spillover?
Findings
5. What finding or result do you see as the study’s most important contribution?
6. Reviewing the specific findings for the thematic areas, based on your expertise, how well did the study answer the evaluation questions regarding: <ul style="list-style-type: none"> a. [list outcome or thematic area] b. [list outcome or thematic area] c. [list outcome or thematic area] d. [list outcome or thematic area]
7. If the study were repeated, what changes would you recommend (if any), and how would these alterations improve the results?
Limitations
8. Are there limitations to the evaluation design, estimation strategy, or data collection approach that are not mentioned, but should be considered and accounted for in the report?
Interpretation
9. Are there conclusions within the context of the evaluation questions and findings presented that are not considered or presented in the evaluation report? What additional analysis would need to occur to draw these conclusions?

Publication
10. While noting that this is not a report prepared for publication, does this evaluation have the potential to meet the standards for peer-reviewed publication?
Additional Comments
11. Do you have additional comments on other important factors?

Resources on Implementation Fidelity Monitoring

While not yet standard practice, it is important to include implementation fidelity monitoring tasks and reports in IE statements of work. These tasks may include regularly scheduled check-in calls between the intervention team, IE team, and donor; monitoring site visits; data transfer and review; or other activities that allow the IE team to track implementation progress. Fidelity monitoring activities can provide critical opportunity for mid-course corrections that align implementation to the IE design or that shift the IE design to align with implementation realities. It can also alert the IE team when potential threats to the validity of the IE need to be addressed (e.g., take-up of the intervention is low, potential contamination of the control group, delays in intervention roll-out).

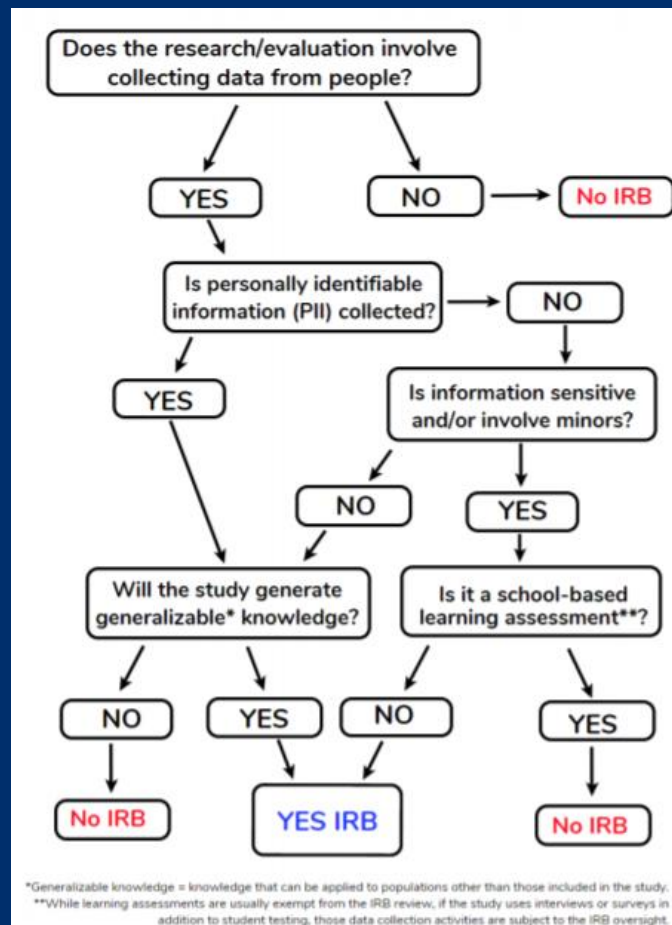
- The USAID/Jordan Monitoring and Evaluation Support Program (MESP) has a useful [presentation](#) on fidelity of implementation for an IE of an app-based literacy library.
- The U.S. Department of Health and Human Services has a helpful [brief](#) on how to measure implementation fidelity, which can provide a sense of the tasks and resources needed.

Resources for Ethics Considerations

- USAID Office of Education: [Policy Brief on Ethics in Research and Evaluation in the Education Sector \(2017\)](#)
- U.S. Global Development Lab: [Considerations for Using Data Responsibly at USAID \(2019\)](#)

BOX 2: ETHICS CONSIDERATIONS IN RESEARCH AND EVALUATION AT USAID

- The Common Rule applies to all USAID-funded evaluation and research projects that involve human subjects. The research or evaluation must also conform to legal and other requirements governing research with human subjects in the country where it is conducted.
- To satisfy USAID CFR requirement of ethics review, a study must be reviewed and approved or deemed exempt by a U.S.-based institutional review board (IRB).
- It is the responsibility of the Contracting Officer's Representative (COR) or Agreement Officer's Representative (AOR) to ensure that all research and evaluation activities funded by USAID comply with USAID's policy on the protection of human subjects as well as all relevant local laws and regulations. It is the responsibility of the award holder to manage the approval process and adhere to study plans approved by IRB and/or local ethics review committee. **Conducting a research or evaluation study without the appropriate ethics review constitutes a violation of USAID's policy on the protection of human subjects.**



Source: *Policy Brief on Ethics in Research and Evaluation in the Education Sector, 2017*

Resources for Cost Analysis

- The USAID Office of Education’s Cost Measurement Initiative has several useful resources for USAID-funded education activities that can also serve as resources for other sectors:
 - [Cost Reporting Guidance \(2018\)](#)
 - [Cost Reporting: Field Implementation Guidance \(2018\)](#)
 - [Cost Analysis Guidance \(2020\)](#)
- USAID is a partner in the Center for Effective Global Action’s [Cost Transparency Initiative](#) and its Costing Community of Practice (CCoP). The CCoP is developing a curated website of cost analysis resources; reporting guidance for researchers and evaluators of cost evidence; and inaugurating a forum to expand access to expertise and build the cost evidence base in the field of international development. Resources from other CCoP members include:
 - Resources from the Abdul Latif Jameel Poverty Action Lab (J-PAL) on [conducting cost-effectiveness analysis](#) include guidance notes, cost collection templates, and a white paper outlining their methodology.
 - The World Bank’s Strategic Impact Evaluation Fund and the International Rescue Committee jointly produced practical [cost capture guidance](#) for collecting rigorous cost data that is disaggregated, intervention-specific, and captured over the course of an intervention.

Resources for IE Language in Contracts/Agreements

Box 3 provides illustrative text for an IP solicitation that can be adapted for contracts and agreements. It should also include language about expected coordination with the evaluator to report cost data for cost-effectiveness analyses, per the new ADS 201 requirement.

BOX 3: ILLUSTRATIVE RFP LANGUAGE FOR IMPLEMENTATION BIDDERS ON AN IE

Request for Task Order Proposals SOL-624-15-00001 for the West Africa Biodiversity and Climate Change (WA-BiCC) Activity for USAID/West Africa (USAID/WA)

“An important component of the WA-BiCC project is to identify which interventions work, and why, to facilitate learning and scaling up across the region. A core objective of the WA-BiCC approach is that regional and national decision-makers use evidence to inform policy and management decisions for biodiversity conservation, climate change mitigation, and adaptation. As a result, USAID/WA will identify opportunities and implement an impact evaluation (IE) over the course of the WA-BiCC project. It is expected that the implementer of the WA-BiCC project implementing partner will not only accommodate, but and proactively support, the implementation of an IE over the life of the project.

Specifically, USAID/WA, in consultation with the USAID/E3 Bureau, shall evaluate approaches to the conservation, restoration, and/or sustainable management of coastal and mangrove ecosystems, and gather evidence to understand the biodiversity, climate change mitigation, and climate adaptation outcomes these interventions provide. USAID/E3 will provide an analysis of some (not all) potential interventions, as well as a potential framework for an IE of some of these interventions. This document is not intended to be prescriptive. However, it is the intention of USAID to rigorously evaluate the impact of the project (or components thereof) within the structure of an experimental or quasi-experimental design. The WA-BiCC contractor will be expected to accommodate modifications to its approaches, implementation schedule, and monitoring framework to facilitate such an evaluation. This may include modifications, including, but not limited to, the selection, timing, location, and way activities are implemented and the monitoring methodology used. It is expected that the contractor will collect certain data (or facilitate the collection of such data by other partners selected by USAID). The purposes served by these data include: the selection of intervention and/or control or comparison sites; to understand baseline conditions and post-intervention outcomes; to understand biophysical and socio-economic variables which that may identify the conditions in under which project approaches are (or are not) successful; or other purposes, as appropriate. USAID/WA expects to count on independent evaluation expertise to advise the project and carry out aspects of the evaluation; however, the WA-BiCC contractor will need to coordinate some aspects of project implementation and monitoring with USAID and the external evaluation contractor. Expectations and coordination will be discussed, agreed to, and articulated during the WA-BiCC annual work planning process.

Given that USAID intends to conduct an IE of certain biodiversity and climate change approaches in coastal areas, the offeror shall identify (within the discussion of their technical approach) those interventions and associated higher-level biodiversity and climate change outcomes (not outputs) which it thinks are particularly suitable for an IE. The technical approach will suggest an implementation and monitoring approach which that the contractor estimates will facilitate the measurement those primary outcomes, as well as any secondary or co-benefits (e.g., socio economic, health, governance, etc.) which that the contractor expects may result from the project. The IE will be implemented under a parallel mechanism that will be separated, arranged, and supervised by USAID.

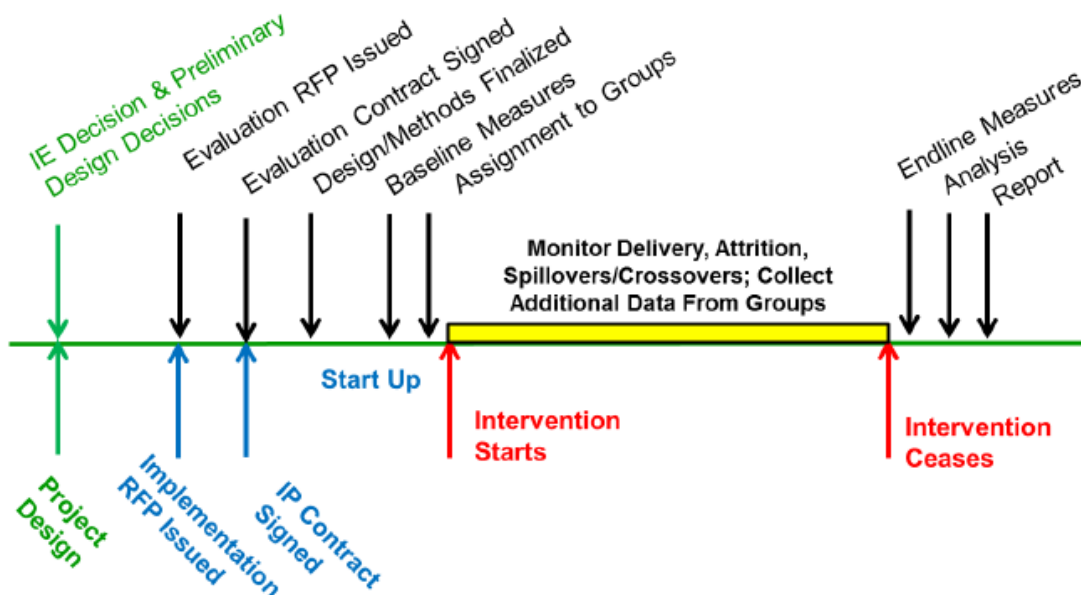
USAID evaluation policy also requires that baseline data from both target group and comparison group be collected prior to the start of the intervention under study. To this end, bidders should expect that upon award, USAID will facilitate discussions between its implementing partner and evaluation team that result in the establishment of a parallel start-up period, during which baseline data will be collected on the activity’s performance indicators and the IE’s outcome variables, and that both baselines will be delivered to USAID before initiation of the activity’s deliverance of the intervention under study is authorized.”

Resources for Coordination and Collaboration

- The E3/Land and Urban's [Impact Evaluation Brief](#) outlines key steps taken during the evaluation process by the IP, the IE team, and USAID to ensure coordination and collaboration throughout the IE process.
- Book: Glennerster, Rachel, and Kudzai Takavarasha. Running randomized evaluations: A practical guide. Princeton University Press, 2013. <http://runningres.com/> (website includes useful resources)

Timing the Issuance of an IE

For an IE to succeed, a partnership needs to be established between the IE team and the team that is commissioned to implement the intervention that will be studied. Given the need for a control or comparison group and multiple rounds of data collection, including a pre-intervention baseline, IEs need to start up before the intervention IP begins to deliver whatever goods or services, or policy reforms, constitute the intervention. When planning an IE, USAID evaluation managers and technical officers need to consider the timing issues illustrated in the graphic below. In some cases, starting in parallel with an IP is critical. In other cases, when the IP is expected to have a long start-up period to research and design an intervention, a later start for the IE team may be possible. What is important is for USAID evaluation managers and technical officers to understand the timeline for intervention start up and initiate evaluation preparations on a timely basis in relation to that timeline.



IE Process Management Checklist

USAID's [Evaluation Toolkit](#) includes an evaluation process management checklist that was designed to help Agency staff responsible for overseeing an evaluation do that job well. The checklist has no features specific to IEs. The discussion paper authors customized USAID's existing checklist to make it a valuable resource for IE managers. Changes in USAID's existing checklist are most noticeable in the IE-specific standards for a range of checklist tasks and milestones. In addition, the main opportunities for quality control over the course of an IE are highlighted.

IMPACT EVALUATION PROCESS MANAGEMENT CHECKLIST

(Adapted from USAID's *Evaluation Process Management Checklist* [2015])

Key: = Quality control opportunity

Task/Milestone	Target Date
Planning	
Preliminary decision made to explore undertaking an impact evaluation	
Evaluability assessment conducted, inclusive of: <ul style="list-style-type: none"> • Stakeholder analysis to determine information needed, priority questions, likelihood of use • Theory of change (developed or refined) <input checked="" type="checkbox"/> • Existing data assembled, including performance indicators and data (pre-evaluation clinics can facilitate this process) 	
Decision made to proceed with impact evaluation	
Evaluation manager/COR or designee named	
Evaluation manager's impact evaluation management work plan developed	
Stakeholder input obtained, as appropriate (if evaluability assessment was not done)	
Main impact evaluation parameters determined (e.g., purpose, questions, timing)	
Mission-wide evaluation plan updated (if appropriate)	
Initial evaluation communication and dissemination plan developed	
Evaluation SOW drafted, including standards for design and final report <input checked="" type="checkbox"/>	
Evaluation Independent Government Cost Estimate drafted	
Peer review of evaluation SOW against impact evaluation-specific SOW outline <input checked="" type="checkbox"/>	
Final evaluation SOW approved	
Mechanism selected (for external evaluations)	
Solicitation issued (for external evaluations)	
Technical evaluation of proposals (review) <input checked="" type="checkbox"/>	
Evaluation contract awarded (external) or evaluation team selected (internal)	
Disclosure of conflict of interest forms received	
Design	
Post-award orientation ("kickoff") meeting with evaluation team <input checked="" type="checkbox"/>	
Evaluation work plan submitted by evaluation team	
Evaluation background review/literature review submitted (if requested)	
Post-award evaluation workshop with evaluation team and activity implementation team (recommended)	
Scoping visit to determine feasibility of design options (if possible); report submitted	
Evaluation design report submitted (based on design report standards/outline included in evaluation SOW)	

Task/Milestone	Target Date
Evaluation design shared with partner country stakeholders and implementing partners	
Internal USAID peer review of impact evaluation design report <input checked="" type="checkbox"/>	
External peer review conducted by impact evaluation expert(s) <input checked="" type="checkbox"/>	
Evaluation design approved (after modifications if required)	
In-brief for mission and/or evaluation stakeholders (if requested)	
Implementation and Reporting	
Stakeholder engagement to select treatment and comparison sites	
Stakeholder engagement during baseline planning activities	
Out-briefing from baseline data collection (if requested)	
Submission of baseline draft report	
Internal USAID peer review of draft report <input checked="" type="checkbox"/>	
Submission of baseline final report	
Acceptance of baseline final report by USAID mission/operating unit	
Stakeholder engagement during implementation fidelity monitoring activities and reports <input checked="" type="checkbox"/>	
Mid-evaluation briefing/periodic briefing (if requested)	
Out-briefing (if requested)	
Stakeholder engagement during endline planning activities	
Out-briefing from endline data collection (if requested)	
Evaluation team's preliminary briefing on evaluation findings, conclusions, recommendations; feedback provided by USAID, other stakeholders <input checked="" type="checkbox"/>	
Submission of draft evaluation report	
Internal USAID peer review of draft report <input checked="" type="checkbox"/>	
External peer review conducted by impact evaluation expert(s) <input checked="" type="checkbox"/>	
Statements of differences by funders, implementers, and/or members of the evaluation team received	
Submission of final evaluation report	
Acceptance of final evaluation report by USAID mission/operating unit	
Dissemination and Utilization	
Evaluation report dissemination plan updated	
Evaluation report disseminated	
Evaluation report submitted to DEC following COR approval	
Evaluation data submitted to the Development Data Library (DDL) (if applicable)	
Evaluation contractor performance assessed (CPAR system) (if applicable)	
USAID post-evaluation review of findings, conclusions, and recommendations	

Task/Milestone	Target Date
Post-evaluation action plan approved	
Actions in post-evaluation action plan completed	
Evaluation summary data entered into the evaluation registry of the performance plan and report (PPR)	

IE MANAGEMENT PLAN

In addition to a checklist of tasks to be completed, a management plan for an IE can include:

- A list of deliverables and standards for the acceptance of those deliverables
- A schedule that shows when tasks will be initiated and completed, and when deliverables will be submitted. Bidders can be asked to include a Gantt chart in their submission and to periodically update it in their quarterly reports, or a USAID evaluation manager can create their own

	Year 1				Year 2				Year 3				Year 4			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Planning																
Task/Milestone																

- A summary of the roles and responsibilities of USAID staff who are members of the internal IE team. A simple way to capture this information is by extending the number of columns shown on the evaluation manager’s checklist. For example:

Planning	Target Date	Evaluation Team Leader (External)	USAID Evaluation Manager	Mission M&E Officer
Task/Milestone				