CAPACITY ASSESSMENT

FOR

RESEARCH, EVALUATION, STATISTICS,

AND OTHER ANALYSIS

ACRONYMS

ADS	Automated Directives System
AOR	Agreement Officer's Representatives
CDCS	Country Development Cooperation Strategy
CLA	Collaborating, Learning, and Adapting
COR	Contracting Officer's Representative
DA	Development Assistance
DDL	Development Data Library
DEC	Development Experience Clearinghouse
DIS	Development Information Solution
DMP	Data Management Plan
DO	Development Objectives
ER	Evaluation Registry
GH	Global Health
MEL	Monitoring Evaluation and Learning
MMBT	Maturity Matrix Benchmarking Tool

NOFO	Notice of Funding Opportunity
OECD	Organization for Economic Co-operation and Development
OMB	Office of Management and Budget
OU	Operating Unit
RSF	Resilience and Food Security
SME	Subject Matter Expert
TOSSD	Total Official Support for Sustainable Development

TABLE OF CONTENTS

1. INTRODUCTION

Evidence generation and use at USAID is decentralized, with central offices in Washington DC providing policy, guidance, and technical assistance to support the operations of Missions in Africa, Asia, Europe and Eurasia, and Latin America and Caribbean. The decentralized structure of the Agency means that evaluations, research, statistics, and analysis efforts, are generally context specific. Decentralization provides opportunities for USAID to generate and use evidence, by leveraging the interests and technical specialization of staff across different geographic areas. USAID's capacity to generate and use evidence is also guided by policies that address evaluation, research, and data. These include the Evaluation Policy of 2011, updated in 2016 and incorporated into ADS 201 (focus on evidence requirements in the Program Cycle), the USAID Scientific Research Policy, and the USAID Open Data Policy.

The Foundations for Evidence-Based Policymaking Act (Pub. L. 115–435) "Evidence Act", requires agencies to submit a Capacity Assessment for Research, Evaluation, Statistics, and Other Analysis every four years as part of their Strategic Plans. Led by the Agency Evaluation Officer, in conjunction with the Statistical Official, Chief Data Officer, and other Agency personnel, agencies are required to conduct and provide an assessment of the (1) coverage, (2) quality, (3) methods, (4) effectiveness, and (5) independence of the statistics, evaluation, research, and analysis efforts of the Agency. As part of the Capacity Assessment for Statistics, Evaluation, Research, and Analysis agencies must address the following:

- List existing activities and operations that are being evaluated or analyzed;
- Determine the extent to which the evaluations, research, and analysis efforts and related activities support the needs of the agency;
- Determine the extent to which the evaluation, research, and analysis efforts and related activities of the agency address an appropriate balance between needs related to organizational learning, ongoing program management, performance management, strategic management, interagency and private sector coordination, internal and external oversight, and accountability;
- Determine the extent to which evaluation and research capacity is present within the agency, including personnel and agency processes for planning and implementing evaluation activities, disseminating best practices and findings, and incorporating employee views and feedback;
- Determine the extent to which the agency has the capacity to assist agency staff and program offices to develop the capacity to use evaluation research and analysis approaches and data in the day-to-day operations; and

- Determine the extent to which the agency uses methods and combinations of methods that are appropriate to agency divisions and the corresponding research questions being addressed, including an appropriate combination of formative and summative evaluation research and analysis approaches.

This report provides an analysis of how these criteria apply to statistics, evaluation, research, and analysis functions and activities within USAID.

2. METHODOLOGY

Data for the Capacity Assessment report was derived from (1) a review and analysis of USAID documents and tools for generating, managing, and using evidence, and (2) a capacity assessment study conducted by a third-party contractor, to assess USAID's level of maturity to generate, manage and use evidence.

Document Review - Documents reviewed included published and unpublished literature produced by the Agency. Among these were evaluation reports, policy documents, research reports and guidelines, and learning documents. Findings from the desk review were used to inform the assessment of the coverage, quality, methods, effectiveness, and independence of evaluation, research, statistics, and analysis across the Agency.

Capacity Assessment Study - The study used a mixed-methods approach, drawing on both qualitative and quantitative data. Data was drawn from key informant interviews, focus group discussions, and an Agency-wide survey of 637 staff members across the Agency. Several workshop discussions were also conducted by the study team to gather diverse perspectives on evidence management and use from Agency staff across USAID Missions, Bureaus, and Independent Offices. Findings from the study were used to develop an Evidence Maturity Matrix that benchmarked current USAID capacity to generate, manage, and use evidence.

2.1. DEVELOPING THE MATURITY MATRIX BENCHMARKING TOOL

The Maturity Matrix Benchmarking Tool (MMBT) is tailor-made to USAID and the Agency's own understanding of capacity in evidence generation, management, and use. Working with a capacity assessment study team, USAID first held a capacity assessment design session to shape the assessment questions, and to draft the assessment framework. Applying the concept of USAID's Collaborating, Learning and Adapting framework (CLA), a capacity assessment framework was developed around the following five domains - (1) resources, (2) culture, (3) collaborating, (4) learning, and (5) adapting - elements that are most critical to USAID evidence generation, management, and use. The team also identified thirteen parameters of capacity across the five capacity assessment domains. Figure 1 below shows the capacity assessment framework.
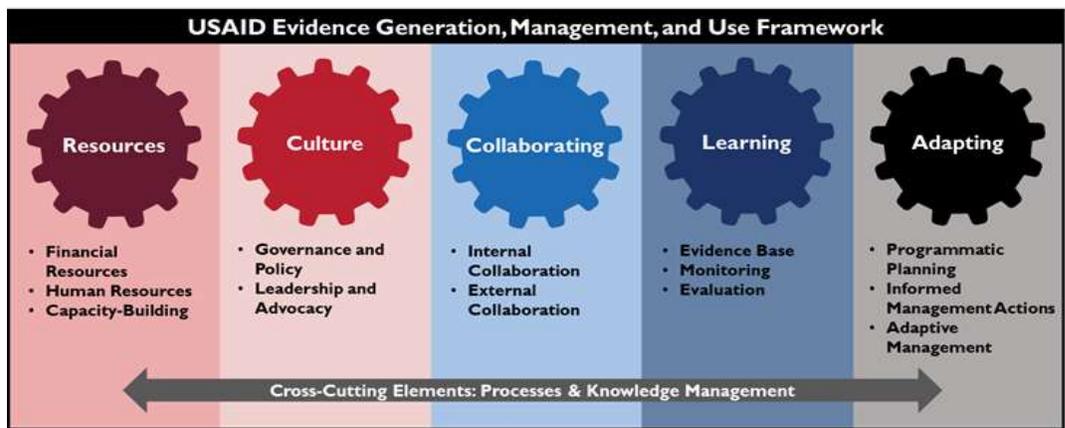
Figure 1 - The Capacity Assessment Framework

This was followed by three design sessions conducted with selected PPL/LER staff, and USAID subject matter experts (SMEs), who then documented and defined what USAID considers to be "limited" and "full" evidence capacity, within the identified parameters of the capacity assessment framework. Using this definition, and the identified domains and parameters of capacity, the MMBT was developed. This tool was then used to assess and create a benchmark of current USAID capacity to generate, use, and manage evidence. For this assessment, which focuses specifically on generation, management, and use of evidence, it was agreed to focus on seven of the thirteen parameters across the five domains of the capacity assessment framework. These are (1) Evidence Base, (2) Programmatic Planning, (3) Adaptive Management, (4) Leadership and Advocacy, (5) Internal Collaboration, (6) External Collaboration, and (7) Capacity Building.

The capacity assessment framework and the maturity model benchmarking tool were used to rate the 'maturity' level of the Agency within each of the five domains. The maturity levels for each domain were defined as (1) Basic, (2) Developing, (3) Advanced, and (4) Leading. The tool was used to diagnose the current state of the organizational capacity to generate, use and manage evidence. Based on the diagnosis, and level of maturity, targets can be set for future improvement in the level of maturity as appropriate.

2.2. DATA COLLECTION FOR THE MATURITY MATRIX

The MMBT was used to support the development of data collection instruments (i.e., interview guides, a survey, and data interpretation workshop materials) which asked targeted questions about staff's experiences in the areas of evidence generation, management, and use, determined to be most important for USAID. A survey was administered to USAID staff in January 2021. A total of 637 staff responded to the survey. The survey targeted USAID Washington, D.C., and Mission staff involved in activity design and/or implementation within the past two

years. The survey was sent to 6,196 staff members, with an explicit "pre-screen" page so that those not involved in activity design and implementation could self-exclude. The survey response rate was 13.5 percent, as 836 staff members responded. However, of that number, only 637 respondents met the criteria of the screening question of having been involved in activity design and/or implementation within the past two years and were included in the analysis.

Among the respondents, 50 percent worked in technical sector roles, 42 percent were in AOR/COR roles, 29 percent were working in Monitoring Evaluation, and Learning (MEL) or data analytics, 24 percent were working in program office processes, 14 percent were in procurement and contracting, and 5 percent were in a Mission front office.

The following questions were developed to guide the assessment of the Agency's current capacity to manage and use evidence:

1. To what extent is the Agency incorporating evidence into its programmatic work?
2. What are the ways in which Agency staff access, as well as facilitate the use of, evidence?
   a. Are there ways in which staff access or facilitate the use of evidence that appear to be more effective than others?
3. What are the Agency's approaches to building staff capacity in evidence management and use?
   b. Are there capacity-building approaches being implemented which appear to be more effective than others?

These questions guided the development of focus group discussion guides, and the survey that was administered. While the questions did not directly discuss evidence generation, the survey included questions on evidence generation. Findings and conclusions addressing these questions are intended to inform action-oriented practical recommendations and guidance to further build capacity across the Agency.

3. THE EVIDENCE MATURITY MATRIX

Using quantitative data from the staff survey, and qualitative data from key informant interviews (KII), focus group discussions (FGDs), and data interpretation workshops, the MMBT was used to develop maturity levels of the Agency, to manage and use evidence, across the seven parameters of the five domains of the assessment framework, as follows:

3.1. RESOURCES

Capacity Building (Maturity Level - Developing): Within the domain of resources, capacity building was assessed as developing. Four elements of capacity-building were assessed. These included (1) availability of capacity-building resources, (2) planning and organization of training, (3) use of resources across Operating Units, and (4) incorporation of Monitoring, Evaluation, and Learning (MEL) activities in work plans and contracts. A developing maturity level in capacity building means that training events are somewhat well planned and organized. Also, staff are provided occasional feedback on training and learning needs regarding evidence generation, management, or use, which are occasionally included in training planning. At this level, capacity-building resources are sometimes used, although often unequally, across operating Units. In addition, capacity-building activities in MEL and CLA are sometimes incorporated into necessary work plans and contracts for USAID staff, partners, and implementing partners.

While capacity-building training options were available to respondents of the staff survey, around a third of respondents (32 percent) indicated in the last two years, they have attended or participated in at least one of the individual training events listed in the survey; 28 percent of survey respondents reported that they had not attended any USAID-provided training on evidence management and use in the past two years. About a third of survey respondents (32 percent) noted that they were not aware of when and where training is offered. In terms of data from the key informant interviews, four key informants and participants from both workshops similarly noted that respective teams across the Agency (both at headquarters and at Missions) need to make staff better aware of available/existing capacity-building opportunities and give staff the flexibility to make use of available training opportunities.

When survey respondents were asked if they felt that the training activities and other resources provided by USAID are sufficient to help them meet the Agency's requirements for evidence use, 35 percent felt that training activities are sufficient. An equal proportion (35 percent) also felt it was not sufficient, while 30 percent indicated that they do not know or were not sure if the training offered was sufficient. Among technical officers, thirty-three percent felt that the training activities and other resources were not sufficient, while 36 percent felt that they were sufficient.

When asked about the gaps and/or weaknesses in capacity building training on evidence generation, management, and use provided by USAID, the most common gaps mentioned by respondents are (1) training contents does not specify how teams should operationalize evidence requirements; and (2) inadequate institutional support or incentive for attending trainings.

3.2. CULTURE

Leadership and Advocacy (Maturity Level - Developing): Within the culture domain of the maturity matrix, leadership and advocacy were assessed. This included, (1) Agency senior leadership support for advocacy, and resourcing for evidence; (2) Operating Unit leadership support, advocacy, and resourcing for evidence; and (3) the extent to which processes are in place to access and use evidence for decision-making. Based on the available quantitative and qualitative data analyzed, leadership and advocacy were rated at a developing maturity level. The developing maturity level suggests that leadership at the Operating Unit level sometimes supports, advocates for, and allocates resources to evidence generation, management, and use. Operating Unit leaders sometimes encourage and/or incentivize evidence-based decision-making, and evidence is sometimes incorporated into communications about decisions. It also suggests that communications from Operating Unit leaders sometimes include references to data and evidence.

Results from the survey indicated that only 36 percent of all respondents say that in their experience, senior Agency leadership regularly support or promote the use of evidence, by advocating for resource investment in evidence generation, management, and use. In Key Informant Interviews (KIIs), some informants felt that Agency senior leadership was very supportive of evidence use as evidenced by public communications, willingness to allocate funding for evidence, and demand for evidence. However, most others felt that Agency leaders were not strong enough advocates for evidence management and use.

On the other hand, there was a much higher percentage of survey respondents who indicated that in their experience, senior leadership in their operating units (OUs) regularly support or promote the use of evidence. Slightly more than half (52 percent) of all survey respondents indicated that in their experience senior leadership in operating units demonstrated support for evidence use, by regularly requesting and expecting evidence to be presented or shared by activity teams. Also, 48% of respondents indicated that OU senior leadership regularly use evidence in decision making, 47% say they frequently reference data and evidence in communications, and 45% say that OU senior leadership regularly create opportunities and processes to regularly receive evidence from findings and conclusions of evaluations, research, and other analysis.

3.3. COLLABORATING

There were two elements of collaboration that were assessed, and for which evidence maturity levels were established. These are internal collaboration and external collaboration.

Internal Collaboration (Maturity Level - Advanced): In establishing a maturity level for internal collaboration, the following elements were assessed: (1) the extent to which platforms and processes were

available for evidence sharing; (2) the existence of knowledge products; and (3) frequent and systematic engagement across Operating Units. Based on the quantitative and qualitative data analyzed, internal collaboration was rated at an advanced level. In the advanced stage of maturity, distinct roles, responsibilities, and rules governing how staff should collaborate on evidence generation, management, and use within the Agency are established, mostly understood by staff, and are often functioning. Technology platforms and processes are available for data and evidence sharing within the Agency. The benchmark also suggests that many knowledge products are available (e.g., leading practices, reports, toolkits, templates, and one-pagers, etc.) and are often distinguished by intended users and/or audience. Additionally, there is frequent engagement between Missions, and Washington, D.C., Bureaus, and among Operating Units regarding evidence gathering and sharing.

When survey respondents were asked about the stakeholders with whom they engaged in the last two years to inform program activity design processes, 82 percent indicated that they engaged with internal stakeholders from within their own operating unit, and 78 percent engaged with internal stakeholders in other operating units. Among this group, the most common stakeholders with whom they reported collaborating during activity design are other team members from their operating units (87%). In addition, 77 percent of survey respondents said they or their team had consulted with USAID SMEs when designing or supporting the design of a USAID program activity. More than half (58%) of survey respondents indicated that they collaborate with USAID M&E specialists from any operating units during activity design. While only about a third (31%) of respondents reported that during activity design, they collaborate with CLA advisors from any operating unit.

External Collaboration (Maturity Level - Developing): External collaboration was assessed based on the following two elements: (1) availability of processes and mechanisms to facilitate access to external researchers/experts, and (2) attendance at conferences and summits to share evidence. Based on findings from the survey, KII and FGDs, the maturity level for external collaboration was rated as developing. This suggests that basic mechanisms are available to facilitate access to external researchers/experts. Some processes are available for external stakeholder engagement, including with host countries, for evidence gathering, sharing, and use. Also, there are occasionally used venues/platforms for sharing data, evidence, and reports with external stakeholders; and some knowledge products (e.g., leading practices, reports, toolkits, templates, one-pagers) are available.

Over two-thirds (70%) of all survey respondents indicated that they have engaged with external researchers or experts, to generate, gather, share, and use evidence. The two most common activities reported by survey respondents in which external experts are engaged are for conducting evaluations (51%) and sector research (47%). However, among survey respondents who did not engage with external experts for evidence

management and use, 27 percent noted that they did not do so because of insufficient budgets, and 21 percent indicated that processes for engaging external experts are too lengthy or complex.

During workshop discussions, participants stressed that partner country governments and other local stakeholders, communities of practice, and working groups are seen as critical partners in evidence generation, collection, and use. Workshop participants also remarked that the Agency is increasingly investing in co-creation approaches, which participants saw as supportive to increasing external collaboration.

3.4. LEARNING

Evidence Base (Maturity Level - Developing): In assessing the level of maturity of the evidence base of the Agency, two elements were considered. These are (1) the availability of evidence to meet the needs of Operating Units, and (2) the existence of formal approaches for tracking evidence. In the assessment, the level of maturity of the evidence base was rated as developing. An evidence base at the developing level suggests that technical evidence exists to meet some Operating Unit needs, and staff are sometimes familiar with it. Also, staff sometimes feel that they have the evidence they need for programmatic planning or decision-making. It also suggests that formal approaches for tracking the existing technical evidence within USAID and externally are sometimes used. In addition, learning agendas are sometimes available, and sometimes used to prioritize evidence generation activities by Operating Units.

The survey suggests that when staff design, or support the design of USAID activities, they use evidence from a variety of sources. The most common sources of evidence reported by respondents are consulting with USAID subject matter experts (77%), review of evaluations, monitoring data, and lessons learned from prior activities (76%), reviewing data from relevant national or multinational datasets (66%), site visits and/or consultations with potential beneficiaries and local stakeholders (62%), review findings from peer-reviewed publications, academic literature, etc. (57%), and conduct new analysis (54%).

However, when asked if there are any gaps in available evidence to adequately support their activity design needs, over 80% of respondents indicated that there have been gaps in the evidence they need. The most common gaps identified by respondents include gaps in statistical data or analysis (43 percent), gaps in evaluation from similar or previous programing (40 percent), gaps in monitoring data or reporting from similar previous programming (39 percent), gaps in academic research (34 percent), and gaps in experiential learning from within the team (51 percent). In addition, interviews with key informants revealed concerns about the quality of evidence. Some key informants requested guidance regarding what constitutes 'quality' evidence. Others requested guidance on what constitutes experiential evidence. While informants noted that experiential learning is important, it was unclear how to gather, document, and incorporate this evidence into evidence tracking

documents and systems, or when and how it is appropriate to incorporate this type of knowledge in programming.

3.5. ADAPTING

Under the adapting domain of the maturity matrix, two elements were assessed. These are programmatic planning and adaptive management.

Programmatic Planning (Maturity Level - Advanced): A maturity level of advanced in programmatic planning suggests that evidence is often used to inform strategic planning, project, and activity design. Also, sources of evidence are often reliable and documented. It also suggests that Operating Unit teams also draw evidence from third-party assessments and/or evaluations to complement operating unit assessments and evaluations. This includes using additional evidence drawn from host government sources, civil society, the private sector, and other donors.

Key informants largely reported that the available evidence was used for strategy and activity design. In discussions of the CDCS development process, several respondents noted that evidence was well-incorporated into design where possible, and evidence was intentionally used to develop result frameworks. Most survey respondents reported using evidence from previous activities, and evidence cited in existing strategies, when designing new activities. Survey data indicated that during programmatic planning 76 percent respondents reviewed evaluations, monitoring data, and lessons learned from previous activities, and 60 percent of respondents reviewed existing strategies and documents during activity design.

Adaptive Management (Maturity Level - Advanced): Four elements of adaptive management are assessed. These included (1) the existence of scheduled processes for systematic review of evidence; (2) the extent to which Agency staff work with stakeholders to identify successes, challenges, etc.; (3) the frequency of staff to raise and document decisions; and (4) the ability of teams and Operating Units to implement decisions for changes in programming. An advanced level of maturity suggests that processes are often in place for systematic review. and there is use of findings from evidence generation activities in programmatic decisions. Pause and Reflect opportunities are often hosted for staff and partners. Operating Units often work with partners to identify successes, challenges, and subjects that warrant further exploration. Where findings and conclusions are raised, they are often aligned to specific programmatic and operational decisions, and decisions are often documented. Additionally, at this maturity level, planned actions are often tracked and implemented. Operating Units often use data to inform decisions on maintaining, adapting, or discontinuing current approaches and often take action to adapt strategy, projects, or activities as appropriate.

Results from the survey indicated that 63 percent of respondents have been involved in acting on findings from evaluations, statistics, research, or other analyses. However, when asked what some of the barriers to the use of evidence from evaluations, statistics, and research, to adapt/modify activities during implementation, almost half (48 percent) indicated limited flexibility and independence in staff decision-making as the barrier, and 42 percent indicated lack of flexibility in contract mechanisms, which makes it difficult to adapt or change programs that may be required because of the evaluation or research evidence.

3.6. PROPOSED EVIDENCE BUILDING ACTION

As a result of the findings outlined above, USAID is proposing the following actions that will contribute to evidence building and capacity to access, use, and manage evidence:

- Establishment of an annual peer-to-peer evidence exchange event that will bring together USAID staff, and Academics, think tanks, bilateral and multilateral donors, implementing partners (NGOs, contractors, etc.), evaluation and research firms, etc, to share and exchange with each other latest evidence and experience.

- Establish a quarterly report for USAID Missions to routinely provide updates on how evidence from evaluation, statistics, research, and other analysis are being applied in decision-making to make a difference in their programs.

- Explore the idea and feasibility of establishing a Global Development Evidence Bank, a public/private joint venture for a context driven search engine for multi-sectoral development evidence of what works, that can be accessed by all development practitioners. Internal discussions on this idea will be initiated in 2022.

- USAID is developing a knowledge management function in the PPL Bureau, and through this will provide more guidance & resources for knowledge management and the development of simplified knowledge products, such as 1–2-page summaries with infographics and data visualizations.

- Work towards developing training that is tailored to specific roles, such as leadership, supervisors, and managers.

These actions are expected to contribute towards addressing the gaps identified by respondents around engagement with external experts for evidence management and use. Bringing external partners together with USAID staff will initiate a process of relationship building that can lead to long-term benefits for evidence use and management.

4. CAPACITY ASSESSMENT FOR EVALUATION

Evaluation is an important source from which evidence is generated for decision making at USAID. As a result, evaluation at USAID is incorporated into all phases of the Program Cycle, including strategic planning for Country Development Cooperation Strategy (CDCS), and activity planning and design. There are a broad range of activities that are implemented at USAID for generating evidence from evaluations. This section of the report is based on data from the USAID Evaluation Registry, and findings from the survey of USAID staff capacity to generate, manage, and use evidence.

The USAID Evaluation Registry is a central data system for generating evidence on evaluations conducted across the Agency. The evaluation registry is the main source for evaluation reporting, including the number of evaluations completed by USAID each year and how evaluations are used, and for reporting to the Agency Performance Plan and Agency Performance Report (APP/APR). Bureaus and operating units (OUs) at USAID are required to report evaluation data on planned, ongoing, and completed evaluations into the evaluation registry annually. Appendix A provides a list of evaluations that are currently "ongoing" evaluations in the evaluation registry.

In addition to the evaluation registry, USAID executes an Operational Excellence Agenda (OEA) that informs continuous improvements of operations. The Bureau for Management implements a range of research studies to improve management operations, which include, and are not limited to, benchmarking studies, business process reviews, data-driven after-action reviews, and cost savings studies. They address major management challenges at the Agency and seek to produce practical, actionable recommendations that responsible offices and staff can implement. Often this will involve answering key management questions, identifying the means to answer them through existing data, or collecting new information, as well as, identifying appropriate processes for analysis and dissemination.

4.1. COVERAGE

Annual analysis of the evaluation registry data generally provides evidence of the coverage, scope and breadth of evaluations conducted across the Agency. Evaluations at USAID are conducted by all operating units situated worldwide.

USAID evaluation requirements include the following:

- At least one evaluation per strategy Intermediate Result (IR);
- At least one evaluation for activities with budget > $20m; and
- An impact evaluation for any new, untested approach,

These requirements allow for most USAID activities to be evaluated. Analysis of the USAID Evaluation Registry showed that between 2013 and 2020, USAID conducted over 1,500 evaluations of its programs worldwide. Among these evaluations, 90 percent were performance evaluations and 10 percent were impact evaluations.

The evaluation registry data also demonstrates that evaluations conducted by operating units cover the range of program areas of USAID foreign assistance investment. Economic growth, health, and democracy, human rights, and governance (DRG), accounted for more than three-quarters of evaluations completed by the Agency.

The Agency has recently developed an "Evaluations at USAID" dashboard, that shows where evaluations are conducted, the type of evaluations conducted, and the reports of those evaluations.

4.2. QUALITY

In terms of quality, evaluations reported in the evaluation registry are conducted following the standards outlined in OMB M-20-12. One hundred of these evaluations were randomly selected and reviewed by the PPL/LER evaluation team in FY 2020 to determine the extent to which they meet quality requirements. There are plans by the evaluation team for doing this review every two years. Among the quality elements are objectivity, integrity, and utility.

1. **Objectivity:** All required external evaluations in the evaluation registry and management assessments conducted by the Bureau for Management are transparently planned, designed, and implemented. In planning evaluations and management assessments, program and technical staff actively engage with implementing partners of activities being evaluated to ensure that relevant stakeholders were able to provide input to the evaluation planning process.

    For example, since 2010, USAID/PPL/LER has been awarding Indefinite Delivery Indefinite Quantity (IDIQ) contracts to multiple contract holders, for providing Agency wide Evaluation services. In this case, effort is made so that all the IDIQ holders are aware of potential Request for Task Order Proposals (RFTOPs) that Missions plan to compete among the IDIQ holders.

    Transparency within the IDIQ context means making available information that will enhance the quality of service they provide to the Agency, Timely information on potential task orders is very important for

the contract holders to be able to plan effectively and submit proposals to compete for Request for Task Order Proposals (RFTOP) from USAID operating units.

To operationalize transparency under the IDIQ management system, the COR develops quarterly procurement updates, documenting the status of potential task orders. This update covers the status of potential task orders from the statement of work to when task orders are awarded. The procurement update is shared with contract holders quarterly and covers the following:

- Statements of work under review by LER;
- Statements of work approved by LER and sent to OU to prepare RFTOP;
- RFTOPs in procurement competition; and
- Task orders awarded.

For each category, the procurement update is organized by the OU, the type of activity, the name of the activity, and the sector of the activity. As small businesses, information provided on the procurement updates, allow them to engage in early planning for activities in which they may be interested in bidding on when RFTOPs are completed by OUs.

2. **Integrity:** Data in the evaluation registry is consistently reviewed and analyzed to determine the extent to which evaluations are adhering to quality standards. Annually, USAID conducts analysis of the evaluations reported in the evaluation registry, to determine the quality and rigor of the evaluations completed during the fiscal year. Impact evaluations are reviewed to determine whether they are designed as actual impact evaluation with a counterfactual that is statistically similar to the treatment/ intervention group. Performance evaluations are reviewed and analyzed to determine the extent to which the methods align with the evaluation questions asked. The findings from the most recent review of performance evaluations suggests that in some cases, the evaluation questions need more rigorous methods for answering them. Based on this finding, the USAID Learning Evaluation and Research team has developed a process for providing direct technical assistance to USAID Missions in designing performance evaluations, impact evaluations, and ex-post evaluations starting with the significant evaluations that are in the FY 2022 Annual Evaluation plan.

The Agency also adopted tools and techniques of well-established continuous improvement methods, such as Lean Six Sigma, and used them to create its own methodology for business process reviews. The Agency also uses standard methodologies for its other management studies, including for after-action

reviews, to reduce variation in assessment quality. The Agency reviews its methodologies to ensure they are consistent with industry- leading practices.

USAID evaluations reported in the evaluation registry are also conducted with ethical standards applied in the process. Data collected from study participants are managed in an ethical manner that complies with the Common Rule and adherence to the Protection of Human Subjects in Research Supported by USAID: A Mandatory Reference for ADS Chapter 200.

3. **Utility:** Evaluations conducted at USAID are relevant and are intended to be used by internal and external stakeholders. Analysis of the evaluation registry indicates that the purpose for conducting most evaluations are conducted for organizational learning, to informing future activity design or funding, accountability, performance management, and measuring achievement of results. The Agency establishes the Operational Excellence Agenda annually, with Agency senior leadership prioritizing a set of management studies based on alignment with strategic and risk management priorities, potential cost savings, and other factors. Studies included on the Operational Excellence Agenda result in improvements that optimize Agency operational performance by streamlining processes and eliminating sludge; mitigating Agency risks of fraud, waste, and abuse; and increasing operational policy effectiveness.

Data from the survey conducted as part of the capacity assessment indicated that when staff design or support the design of USAID activities, they use evidence from a variety of sources. The most common ways of gathering evidence as reported by respondents are consulting with USAID subject matter experts (77%), and reviewing evaluations, analysis, monitoring data, and lessons learned from prior activities (76%).

However, when asked if there are gaps in the evidence that are used to support the design of USAID activities, the most common gaps identified by respondents were access to statistical data or analysis (43 percent), and access to evaluation from similar or previous programming (40 percent). Based on the findings of the survey, the Agency is working on enhancing processes that will increase access to evaluation results, to increase utility and reduce gaps. For example, plans are being made to develop a quarterly report titled "Evidence from the Field" that will allow for peer-to-peer exchanges of how evidence generated from evaluations and other sources are applied in programmatic decision making.

4.3. METHODS

What are the methods being used for these activities, do these methods incorporate the necessary level of rigor, and are those methods appropriate for the activities to which they are being applied?

USAID evaluations are conducted with the most appropriate design and methodologies to answer key questions in order to generate the highest-quality, and most credible data and evidence that corresponds to the questions being asked - taking into consideration time, budget, scale, feasibility and other practical and contextual issues associated with the evaluation. Also, USAID evaluations produce well-documented findings that are verifiable, reproducible, and that stakeholders can confidently rely upon, while providing clear explanations of limitations, and identify areas and topics for further analyses.

USAID evaluations are either performance evaluations or impact evaluations. Performance evaluations encompass a broad range of evaluation methods. They often incorporate before- and- after comparisons but generally lack a rigorously defined counterfactual. Following definitions in OMB Memo M-20-12, performance evaluations include evaluation types such as, Developmental Evaluation, Formative Evaluation, Outcome Evaluation, Process or Implementation Evaluation.

Impact evaluations (IEs) measure changes in development outcomes that are attributable to a defined intervention, program, policy, or organization. USAID Evaluation Policy requires that each Mission and Washington OU must conduct an impact evaluation, if feasible, of any new, untested approach that is anticipated to be expanded in scale or scope through U.S. Government foreign assistance or other funding sources (i.e., a pilot intervention).

USAID regularly conducts analysis of the evaluations reported in the evaluation registry, to determine the quality and rigor of the evaluations completed during each fiscal year. These include analysis of the type of evaluations conducted, the evaluation methods used, the quality of the evaluation reports, and the type of primary data collection methods used. The results of these analyses are shared with monitoring and evaluation Point of Contacts (POCs) and used to inform improvements in designing and planning evaluations, as well as, to inform the development of additional guidance to improve the quality of evaluations.

USAID recently completed a review of the methods of 133 USAID-funded Impact Evaluations that were completed between FYs 2012 and FY 2019. One of the elements of the assessment was whether the reports met USAID's definition that IEs "require a credible and rigorously defined counterfactual." This determination requires assessing (1) whether the IE report described a control/comparison group, and (2) whether the IE report provided a statistical justification that the control/comparison group is a good comparison for the treatment group. Based on this definition, 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.

In assessing management operations, two of the most common management studies USAID conducts are after-action reviews (AARs) and business process reviews (BPRs). An AAR is an assessment conducted after a project or major activity that allows team members and leaders to discover (learn) what happened and why, reassess direction, and review both successes and challenges. The AAR also identifies next steps or action items toward meeting goals or future tasks after the activity itself has finished. At USAID, a BPR is a systematic approach to improving processes using action research methods to achieve results more effectively and efficiently. BPRs identify ways to eliminate waste that slow down and complicate processes. BPRs go beyond identification of issues and lay out recommendations for how to implement improvements and measure both progress and performance excellence after the BPR concludes.

To improve the rigor of impact evaluations, USAID is updating mandatory references in operational policy guidance documents to ensure that impact evaluation reports clearly show that the chosen comparison group is statistically identical on average with the treatment group at baseline. That is, to demonstrate that if the two groups are identical, with the sole exception that one group participates in the program and the other does not, then any difference in outcome is likely to be due to the program intervention.

In addition, the updated mandatory references require a statistical output table showing the difference between the means of the treatment and control groups with statistical significance. Also, reports will be asked to show the minimum detectable effect for each outcome variable analyzed and qualitative information about behaviors, beliefs and practices that may influence uptake of interventions and outcomes reached. Together, this information will allow improved interpretation of results and implications for future programming.

USAID is also increasing the provision of direct assistance to field Missions in planning, designing, procuring, and managing evaluations. Expert advice will be focused on both impact and performance evaluations having fewer, more concise, and more precise questions. In addition USAID is currently finalizing the development of online and self-paced monitoring and evaluation (M&E) courses that offer learning flexibility and will be easily accessible by staff.

USAID will continue doing annual quality reviews of all evaluations. In addition, given that evaluations are conducted within specific country contexts and field operational settings, USAID will continue to track confounding and/or environmental factors that influence program outcomes.

4.4. EFFECTIVENESS

Are the activities meeting their intended outcomes, including serving the needs of stakeholders and being disseminated?

USAID has an Agency-wide Learning Agenda (ALA). The ALA prioritizes evidence needs related to the Agency's mission to foster development, which covers all development program/sector areas, humanitarian assistance, resilience, and Agency operations. This mission is being articulated in the new USAID's Policy Framework and orienting the Agency's programs, operations, and workforce around the vision of the Biden-Harris Administration. Managed by the Bureau for Policy, Planning and Learning (PPL) the ALA enables the collection, synthesis, and dissemination of evidence, and supports its utilization in learning, to inform USAID's efforts to assist countries to reach better development outcomes.

A data field in the evaluation registry that USAID operating units must complete when reporting on their evaluations is the "evaluation purpose" field. The evaluation purpose field has 10 categories that users can select from. The following list provides the multiple categories that users can select from to describe the purpose for conducting an evaluation that is reported in the evaluation registry:

- **Accountability** - to establish clear responsibility and expectations related to achieving formally approved results;
- **Coordination -** to generate evidence related to the coordination of activities during project implementation;
- **Inform Future Intervention Design or Funding -** to generate evidence to inform decision on future intervention design or funding;
- **Inform Implementation Corrections -** to generate evidence to inform decision making on implementation corrections;
- **Measure Achievement of Results -** to determine the extent to which the intervention has achieved or is achieving the intended and measurable change in the condition it was designed to address;
- **Organizational Learning -** to contribute to knowledge creation and diffusion;
- **Oversight -** to inform the management and oversight of the intervention;
- **Performance Management -** to generate evidence on the achievements of program operations; and track progress toward planned results;
- **Program Management -** to produce evidence for strengthening program operations and improving outcomes; and
- **Strategic Management -** to inform strategy.

Data from this field have been very useful and used primarily to respond to questions that are received from Missions who usually contact the Office of Learning, Evaluation, and Research, asking for examples of evaluations that were conducted for a specific purpose, such as those listed above. In most cases, they use the evaluations

to get a sense of how the evaluation questions relate to the evaluation purpose, in order to inform future design of evaluations.

4.5. INDEPENDENCE

To what extent are the activities being carried out free from bias and inappropriate influence?

All USAID required evaluations are external evaluations commissioned by USAID Missions and Operating Units (OUs) and conducted by an external team external to the Agency and the activity implementing partners.

The required evaluations in the evaluation registry are external evaluations commissioned by USAID Missions and Operating Units (OU), and conducted by an external team, rather than by the implementing partner. In addition, OUs regularly advertise and make a call for Request for Proposals (RFPs) from external evaluators using the SAM.gov website. In addition, all external evaluation team members are required to sign a statement attesting to a lack of conflict of interest or describing an existing conflict of interest relative to the project or activity being evaluated. External evaluators are also required to attest that they have no fiduciary relationship with the implementing partner whose activity is being evaluated.

The Bureau for Management's Performance Team serves as an independent and objective evaluator, leading analysis of management operations in which they are not otherwise materially involved. Analysts utilize business analysis tools widely accepted in the public and private sectors to help the Agency realize continuous performance improvement. Tools used include but are not limited to the following: SWOT analysis, PEST (political, economic, social and technological) analysis, problem tree analysis, business process mapping, value stream analysis, trade off analysis, and cost-effectiveness analysis.

5. CAPACITY ASSESSMENT FOR RESEARCH

To grow its evidence base, USAID invests in basic research, applied research, and experimental development, as defined by the National Center for Science and Engineering Statistics. As a general guideline, scientific research is hypothesis-driven, testable, independently replicable, and quality-controlled through peer-review. It is systematic investigation designed to contribute to generalizable knowledge. To strengthen the generation and use of scientific research evidence to inform programs and policies at USAID, the Agency established a Scientific Research Policy and a Scientific Integrity Policy. In the process of developing these policies, the Agency

conducted extensive consultations with stakeholders and collected data plus other relevant information to identify gaps and develop solutions. The implementation of the Scientific Research Policy is also informed by evidence collected through intra-agency landscape analysis and a needs assessment study.

5.1. COVERAGE

What is happening and where is it happening?

USAID's Scientific Research Policy has instituted guidance for the conduct, management, and use of research. Its five primary components include:

1. Setting research priorities and designing research activities;
2. Quality control standards;
3. Ethical standards;
4. Tracking and reporting on research activities (investments as well as outputs); and
5. Strengthening the scientific and technical expertise within the Agency.

The Agency has instituted reporting requirements through annual performance reporting on research-related programs, and collects indicator data on research outputs, such as peer-reviewed publications, as a result of USAID-funded research projects. This information is used to keep track of planned investments in research development activities and to develop targeted engagements.

USAID's Public Access Plan requires that activities have an approved Data Management Plan and that the data and publications produced are deposited into USAID's repositories (or an approved third-party repository) for public access. USAID is also improving its knowledge management systems to provide easier access to results from USAID-funded scientific and technical activities. For instance, this past year, the Agency created and uploaded to its permanent repository a database of over 21,000 research and technical publications that were supported by USAID but that had not been previously deposited into the repository. This has been accessed approximately 4000 times within the first year of implementation (see below).

5.2. QUALITY

Are the data used of high quality with respect to utility, objectivity, and integrity?

To ensure that research evidence generated and used for making informed decisions is of high quality, USAID's Scientific Research Policy explicitly addresses quality standards in research. Quality controls on a research

activity depend on pre- and post-award actions, and the responsibility rests with the USAID Agreement Officer's Representatives (AORs)/ Contracting Officer's Representatives (CORs) who work with the technical teams designing and managing the study or program, and with the implementing partners. These quality controls include guidance on pre-award actions, such as drafting a well-designed notice of funding opportunity (NOFO), as the first step in ensuring a quality proposal or application. The NOFOs for projects designed to conduct research should provide a clear outline of how the application or proposal should be structured, what content to include, and how the submission will be evaluated. A well written NOFO helps ensure the design of robust research plans that are subject to review by subject matter experts prior to funding decisions.

The final quality of research evidence is achieved through post-award actions. These include periodic review of the work plan, technical review of final research products, as well as peer-review of publications that result from the work. Agency guidance has issued directives that the AOR/COR must ensure that final research reports are reviewed and substantiated for quality before they are disseminated and uploaded to the Agency's permanent archive/repository. In addition, in accordance with OMB guidance, the Agency directives also require that documents deemed Influential Scientific Information or Highly Influential Scientific Assessments must undergo scientific peer review.

5.3. METHODS

What are the methods being used for these activities, do these methods incorporate the necessary level of rigor, and are those methods appropriate for the activities to which they are being applied?

Agency policies direct operating units at USAID to provide implementing partners with clear guidelines outlining the requirements of the research plans and data management plans (DMP) that must be submitted to USAID. All research plans must be peer reviewed by subject matter experts. Scientific peer review is central to the integrity of research and is an accepted standard practice for the U.S. Government (USG) agencies that fund and conduct research. Scientific peer reviewers must be subject matter experts (SMEs) who have in-depth expertise on the topic of the research and should not have a personal, professional, ideological, financial or any other conflict of interest. The reviewers are usually active external researchers in the subject matter of the research and, therefore, qualified "peers" of the primary investigators of the proposed research.

5.4. EFFECTIVENESS

Are the activities meeting their intended outcomes, including serving the needs of stakeholders and being disseminated?

USAID uses scientific research to strengthen the evidence-base of its technical sectors, understand the complex challenges related to international development, develop innovative solutions to those challenges, and bring those solutions to scale. USAID is developing guidance to incorporate research more effectively into programming. The first principle of USAID's Program Cycle Operational Policy is to apply analytic rigor to support evidence-based decision-making. While there are many ways to conduct analysis, scientific research is one of the most rigorous forms. By intentionally incorporating research activities into its work and broader learning initiatives, USAID staff can design more effective programs and see a greater return on investment. Learning efforts at the strategy, program, and activity levels reinforce each other to ensure alignment. Research can be incorporated at the strategy level, during program and activity design and implementation, and during impact evaluations. To do this, USAID Missions, Bureaus, and other operating units must first establish their learning and research priorities. Guidance and resources for establishing these priorities are housed on the USAID Intranet (ProgramNet) pages readily accessible to USAID staff worldwide. In addition, guidance is also incorporated into Agency's Automated Directive System chapters that cover specific issues (e.g., ADS 201 on program design; 578 on peer-review and information quality; 579 on data; etc.).

Research priority setting is an analytical and collaborative exercise for an operating unit's technical, operational, and leadership teams. It is a roadmap or framework that guides inquiry to identify knowledge gaps in Mission programs or Development Objectives (DOs). It provides direction for integrating research into existing programs and for developing new research activities to fill critical knowledge gaps with clear links to higher-level results. This ensures the sustainability of Mission programs and strengthens their strategies.

The following is a list of key considerations for identifying and prioritizing research activities that are included in the ProgramNet resources:

- **Consider the end user:** Before designing any activities, think about who will use the findings

- **Define the problem:** It is often the case that initial impressions about a development challenge will change as more knowledge is gained, leading to deeper understanding and a refined problem statement

- **Identify gaps in knowledge:** What is already known about a given problem? What assumptions need to be tested more rigorously? What is not known with any degree of certainty? Convening subject matter experts, other donors, government officials, local residents, implementing partners, and other key stakeholders for workshops can provide new perspectives and shed light on areas that could benefit from systematic research

- **Question assumptions:** An important part of research priority setting involves being willing to challenge accepted views and assumptions. Repeatedly asking how and why a given situation exists can

lead to new ways of understanding and approaching problems. Questioning the assumptions underlying Theories of Change can also generate important research questions

- **Formulate and prioritize key research questions**: As research topics are being proposed, Mission, Bureau, and office staff should carefully consider how well the topics align with learning agenda questions, as well as with the Mission's collaborating, learning, and adapting (CLA) Plan and MEL Plans.

To improve the effectiveness of implementing USAID's research-related policies, in 2018, a landscape assessment of data and training needs for research was conducted. A follow-on survey of USAID staff was conducted in 2020 to better understand the challenges staff face in their research-related work. The combined results of these studies identified the key needs of USAID staff. Grouped into three broad categories of Using, Managing, and Supporting research, staff indicated needing assistance with the following:

1. USING research evidence in programming and policies
   a. Understanding what scientific research and research evidence are
   b. Finding relevant research evidence
   c. Understanding research findings and how to use them well
   d. Sharing USAID-funded research findings broadly

2. MANAGING USAID-funded research activities
   a. Procuring scientific research
   b. Ensuring research is of good quality
   c. Disseminating research findings for use
   d. Leveraging existing resources

3. Increasing Agency-wide SUPPORT for research
   a. Changing the culture at USAID to value research
   b. Improving the flexibility, time, and money needed for staff to successfully generate, access, and use research
   c. Alleviating USAID processes that impede research

As a result of the findings from both the landscape assessment and the survey, the Agency continues to develop tools and resources that are tailored to the needs of staff and awardees, such as new training modules on the use of research evidence, including impact evaluations.

In addition to using evidence to formulate Agency policies and their implementation, USAID has also made significant efforts in providing access to the evidence generated through USAID-funded research. The Agency's Public Access Plan requires that all products from R&D activities be made publicly available via its Development Experience Clearinghouse (DEC) and the Development Data Library (DDL), the Agency's permanent archive for documents, data and digital objects. The Agency has revised its data policy, found in Chapter 579 of the Automated Directives System (ADS), USAID Development Data, to require Data Management Plans (DMPs) from all programmatic acquisitions and awards that require a Monitoring Evaluation and Learning (MEL) Plan. DMPs will bolster public access to Agency-funded research results and incorporate additional measures to ensure compliance with the Foundations for Evidence-Based Policymaking Act.

A publicly available collection of research articles related to the COVID-19 response entitled, "Pandemic Response Collection in 2020" provides an example of the effectiveness of USAID's Public Access and Scientific Research policies. This collection of over 250 items is available on the Agency's Development Experience Clearinghouse (DEC) and reflects USAID's many years of experience dealing with health crises around the world. Additionally, to further promote dissemination of USAID-funded research products, the Agency plans to join the National Institutes of Health (NIH)-hosted World RePORT -- an open-access, interactive mapping database that highlights biomedical research investments and partnerships from some of the world's largest funding organizations.

5.5. INDEPENDENCE

To what extent are the activities being carried out free from bias and inappropriate influence?

As described above, USAID has drafted Scientific Research and Scientific Integrity policies and guidance that have measures for ensuring the rigor and integrity of USAID's scientific information and research activities. These include guidance on quality control standards, peer-review, and safeguards to prevent bias and inappropriate influence that may compromise the integrity and independence of the information and evidence. USAID is also part of the White House-led Scientific Integrity Task Force and will implement the recommendations that will emerge from that body.

6. CAPACITY ASSESSMENT FOR STATISTICS AND OTHER ANALYSIS

While USAID has no official Statistical Units as defined by the Evidence Act and other pieces of legislation, there are strong decentralized and often siloed statistical programs and analytic capabilities throughout the Agency.

6.1. COVERAGE

What is happening and where is it happening?

As the primary aid Agency for the United States, USAID collects and publishes Development Assistance (DA) data and statistics on behalf of the U.S. Government and has many tools, including Foreign Aid Explorer, that are widely used as the main reference for U.S. development assistance data. USAID also collects and reports U.S. Official Development Assistance figures to the Organization for Economic Co-operation and Development (OECD) Development Advisory Council. Worldwide, USAID's Missions conduct rigorous statistical analyses as part of Monitoring, Evaluation, and Learning Plans and activities. Beyond these overarching efforts, several Bureaus have particularly robust statistical or analytic programs:

- **The Office of the CIO in the Bureau for Management (M/CIO)** handles and synthesizes large amounts of data in order to produce various reports and drive websites. M/CIO houses many statisticians and data scientists that use statistical methodology to draw conclusions, identify trends, and aid programmatic decision-making;

- **The Bureau of Policy, Planning and Learning (PPL)** coordinates with the Office of Management and Budget to represent U.S. interests in international statistical fora such as OECD/ Total Official Support for Sustainable Development (TOSSD) and the UN Statistical Commission; and

- **Several USAID technical bureaus, including Global Health and Resilience and Food Security (RFS),** collect data and conduct statistical analyses to understand trends and formulate responsive programs. For example, the Famine Early Warning Systems Network (FEWSNET) analyzes data and information to provide an outlook on acute food insecurity around the world.

Nonetheless, staff surveyed for this effort did report gaps in the use of statistics and analysis. Forty- three percent of survey respondents reported one or more gaps in statistical data or analyses available to adequately support activity design in the past two years. In line with the most expressed gap in statistical data and analyses, informants reported that the reliability and quality of host country statistics is variable – as one informant stated, "It's very questionable evidence that they base their decisions on given country statistics". Several key informants have also reported that sometimes the evidence base is robust at one level (for example, at a national or meso-level) but not at a higher (for example, the regional level) or a lower level (for example, the

local level). One informant from a Pillar Bureau emphasized that data are sometimes "Not at the appropriate, usable scale to inform research."

6.2. QUALITY

Are the data used of high quality with respect to utility, objectivity, and integrity?

ADS 201 Operational Policy for the Program Cycle requires that OUs conduct Data Quality Assessments (DQAs) for all performance indicators that they report. To ensure performance- indicator data are credible and sufficient for decision-making, OUs and implementing partners are required to uphold five standards of data quality--validity, integrity, precision, reliability, and timeliness. When data do not meet one or more of these standards, OUs are expected to document the limitations and establish plans for addressing them. However, survey respondents from the Capacity Assessment also reported that, "Often the AORs and CORs who are overseeing evaluation activities are not evaluation experts or do not have the right level of knowledge." As such, it is likely that not all AORs/CORs are able to thoroughly assess DQAs and confirm that the data being produced are of a high quality.

However, based on data collected by the Management Bureau, via a random sample of 30% of USAID staff (18% response rate, N = 680) surveyed in the summer of 2020 for an assessment of USAID's workforce data literacy and skills, USAID observed that the majority (nearly 80%) of USAID staff who regularly use data confirm that data are high quality before analyzing them, either by examining the metadata and documentation, examining the data source, or examining the data themselves. Therefore, though it is likely that some of the data being produced are low quality, most USAID staff who regularly use data, assess data quality before use. As a result, USAID staff likely avoid using low quality data for statistical purposes or other analyses for decision making purposes.

6.3. METHODS

What are the methods being used for these activities, do these methods incorporate the necessary level of rigor, and are those methods appropriate for the activities to which they are being applied?

Numerous methods are used for statistical and analysis efforts across USAID. These include but are likely not limited to descriptive and exploratory statistics, visual analysis (e.g. Tableau), geospatial analysis (e.g. using ArcGIS or QGIS), standard regression approaches (e.g. OLS), advanced regression approaches (e.g. fixed and random effects, difference-in-differences), and Artificial Intelligence and Machine Learning approaches. Qualitative evidence suggests that these methods likely incorporate the necessary level of rigor in some cases

and in other cases do not, depending on workforce capacity. The same holds true for whether the methods are appropriate for the activities to which they are being applied. One interviewed Mission respondent reported that, "only some staff really have a strong ability to analyze data. A data specialist in the Mission is very valuable to help people where there is difficulty." Multiple informants noted that there were few internal staff truly dedicated to focusing on data management and analysis.

6.4. EFFECTIVENESS

Are the activities meeting their intended outcomes, including serving the needs of stakeholders and being disseminated?

Though it is challenging to assess whether statistical and analysis activities are meeting their intended outcomes, USAID does maintain a number of platforms for the dissemination of data and analyses to stakeholders. These include repositories for USAID funded data and reports (Development Experience Clearinghouse, Development Data Library) as well as platforms that provide access to third party data and analyses (e.g. International Data & Economic analysis (IDEA)). USAID also works daily with overseas Missions to help them better utilize their data for decision- making through ad-hoc statistical, descriptive, and geospatial analyses.

6.5. INDEPENDENCE

To what extent are the activities being carried out free from bias and inappropriate influence?

To promote statistics and analyses that are free from bias and inappropriate influence, USAID maintains a Scientific Integrity Policy. This policy outlines the principles the Agency is expected to follow to ensure the integrity of its scientific and scholarly activities, including how they are supported and carried out, and research findings are used and disseminated. These principles pertain to two relevant aspects of the Agency's activities for the purposes of this assessment: protecting the scientific process from misconduct and from inappropriate influence and ensuring quality, methodological rigor, and ethical standards in all USAID-funded research activities.

7. CAPACITY BUILDING EFFORTS

OMB Circular No. A-11 states that, "agencies must also address the extent to which the agency has the capacity to assist agency staff and program offices to develop the capacity to use evaluation research and analysis approaches and data in the day-to-day operations."

Based on the Capacity Assessment survey, data was collected on USAID approaches to building capacity in evidence management and use. There were four elements of capacity building for which data was collected. These include: (1) availability of capacity-building resources; (2) planning and organization of training; (3) use of resources across Operating Units; and (4) incorporation of MEL activities in work plans and contracts.

Findings from the survey of staff, conducted for this capacity assessment, indicated that over the last two years, more than a third of the 637 survey respondents participated in an Agency webinar on evidence generation, management, and use. When respondents were asked if they felt that the training activities and other resources provided by USAID are sufficient to help them meet the Agency's requirements for evidence use, 35 percent felt that training activities are sufficient. An equal proportion (35 percent) indicated that training was not sufficient, while 30 percent indicated that they did not know or were not sure if the training offered was sufficient. Among technical officers, 33 percent felt that the training activities and other resources were not sufficient, while 36 percent felt that they were sufficient. To fill this gap, USAID is currently finalizing the development of online and self-paced monitoring and evaluation (M&E) courses that offer learning flexibility and will be easily available to staff.

When asked about the gaps and/or weaknesses in capacity building training on evidence generation, management, and use provided by USAID, the most common gaps mentioned by respondents are: (1) that training contents do not specify how teams should operationalize evidence requirements; and (2) that institutional support or incentive for attending trainings is inadequate.

Respondents were asked in the survey what they would find helpful, to support them or their team's ability to generate, manage, and use evidence. In response to this question, 60 percent of all respondents indicated that simplified knowledge products, such as 1–2-page summaries with infographics and data visualizations will be most helpful. This was followed by tools for managing evidence and knowledge management, identified by 48 percent of respondents; and 41 percent of respondents who say better access to evidence, such as journals, and sector summaries of evidence, will be helpful support for them.

7.1. EVALUATION

Training has always been central to USAID's capacity building efforts. For evaluation capacity building, there are two online evaluation training courses that are offered to staff, through the USAID University. These are the

"Monitoring & Evaluation Essentials: Understanding the Basics", and "Planning and Managing Evaluations" online courses. The focus of these online evaluation courses is to enhance the knowledge of USAID staff on the critical elements of planning and managing evaluations—from learning when one should evaluate, to using the findings from the evaluation. It is expected that after completing these courses, USAID staff will be able to use the knowledge gained, for developing evaluation Statements of Work, managing evaluations, and preparing for the pre-and post-procurement phases of an evaluation. Based on data from the 2019 Annual Performance Plan and Report (APPR) indicator, "percent of USAID staff who have received evaluation training," 32% or 3205 USAID staff received training in evaluation. (*It's important to note that not all USAID staff carry out evaluations - however, it is expected that staff should be familiar with the importance of evaluations and evidence-based decision making. USAID will explore developing training that is tailored to specific roles, such as leadership, supervisors, managers, etc.*)

In addition to the online courses, the decentralized nature of USAID's evaluation functions has allowed the Agency to organize evidence capacity building activities, at the regional level. Across the Agency, regional workshops and clinics are organized by USAID Missions in collaboration with Washington Bureaus, to provide Agency staff with opportunities to develop skills for day-to-day implementation. For example, in the last two years, regional workshops on monitoring, evaluation and learning (MEL) have been sponsored by Africa, Asia, and the Latin America and Caribbean (LAC) Bureaus. These workshops bring together MEL and program staff from Missions in the region, to explore the practical application of evaluation principles and innovative approaches in their day-to-day work. The evaluation clinic held in the Africa region in FY 2018, brought together 52 participants from 26 operating units; similarly, a workshop in the Asia region brought together 27 participants from 14 operating units.

The workshops are designed to be participatory, covering topics with direct application in the day-to-day operations of the work that participants do. Some of the topics that have been covered include, MEL planning; building a Theory of Change Logic Model; writing evaluation Statements of Work and evaluation questions; evaluability assessment, and how to include stakeholders throughout the evaluation process.

At the end of USAID LAC activity, the MEL team asked participants 3 things they learned, 2 things they wanted to learn more about, and 1 action they planned to take. The participants have formed a MEL network through which they continue to do peer-to-peer learning and get support from the regional bureaus in Washington. The Africa Bureau is planning to host another regional clinic in 2022, and in planning for this clinic, the bureau will send out a survey including a question on how participants of past clinics used the skills gained and the extent to which what they learned is contributing to improving evaluation practices in their Mission.

The Agency has also produced the [Evaluation Toolkit](#), which is a source of guidance, tools and templates for initiating, planning, designing, managing, and learning from evaluations at USAID. It is housed on the USAID Learning Lab page and serves as a resource for USAID staff. There are 47 resources contained within the Evaluation Toolkit on Learning Lab. Based on the Learning Lab statistics, from the period October 2017-August 2021, the combined unique web page views for all resources in the toolkit was 107,210. This is in addition to the total unique page views for the landing pages in the Evaluation Toolkit, which totaled 102,765 viewers for the same time period.

To build capacity to support continuous process improvement, the Bureau for Management offers training courses and workshops which promote a standard approach to strategic operations management and business analytics across the Agency. These training courses and workshops help participants learn and apply analytic business tools to real-life operations issues and catalyze an Office or Mission's efforts to strategically plan and launch an operations performance management system.

7.2. RESEARCH

The USAID Scientific Research Policy explicitly addresses the issues of strengthening scientific and technical excellence within the Agency. To address the identified challenges and build capacity of internal and external stakeholders, the Agency conducted a landscape analysis of available training resources offered by various Operating Units at USAID and external partners/organizations. These have been categorized into the following three categories – (1) Learning Opportunities on Planning and Designing Research, (2) Training on Conducting Research, (3) Learning Opportunities on Dissemination, Publication, and Utilization of Research- and can easily be accessed by staff through the USAID ProgramNet intranet site.

In addition to the resources on ProgramNet noted above, the Research Division of the Innovation, Technology, and Research Hub at USAID also developed the training resources listed in the table below for USAID staff. The training on Use of Research Evidence was piloted in person in 2019, was adapted to a blended virtual and online course in 2020 and is now modified and implemented in collaboration with interested operating units across the Agency on a case-by-case basis. For example, in 2021, the Research Division worked with the Bureau of Policy, Planning, and Learning to offer the course to 30 participants across geographies, sectors, and technical backgrounds; is working with two sector-specific operating units to offer a more audience-targeted course to 30 participants each in the democracy and governance as well as the biodiversity sectors; and is discussing working with a geographical bureau to offer the course to a leadership audience. Participation is open to USAID staff of all levels and hiring mechanisms.

TRAINING RESOURCES FOR RESEARCH

Using Research Evidence to Accelerate International Development Outcomes

IN-PERSON TRAINING FOR USAID STAFF
Includes final slide decks, handouts, activity guides, and facilitator guides

MODULE 1: Research Evidence and the USAID Program Cycle
The goal of this module is to help participants understand the value of research evidence and learn how it can be used to inform decision-making throughout the USAID Program Cycle.

MODULE 2: Finding, Evaluating, and Synthesizing Research Evidence for Decision-Making
The goal of this module is to build participants' skills in finding, evaluating, and synthesizing research evidence from individual studies and broader bodies of research to inform decision-making.

Communicating Research Evidence to Inform Policies and Programs

IN-PERSON TRAINING FOR USAID STAFF
Includes a final slide deck, handouts, and facilitator guide.

The goal of this training is to build participants' skills in communicating research findings in concise and action-oriented ways that can be understood by a variety of non-technical audiences, including USAID decision- makers.

Using Research Evidence at USAID

ONLINE TRAINING FOR USAID STAFF ON USAID UNIVERSITY

The course is available to be taken at will on the USAID University website. This is a self-paced training that should take approximately 1.5-2 hours to complete. The goal of this online learning resource is to help participants understand the value of research evidence and learn how it can be used to inform decision-making throughout the USAID Program Cycle. The resource will also build participants' skills in finding, evaluating, and synthesizing research evidence from individual studies and broader bodies of research to inform decision-making.

Using Research Evidence in [Sector or Geography]

BLENDED VIRTUAL & ONLINE TRAINING FOR USAID STAFF
Includes final slide decks, handouts, activity guides, and facilitator guides

The first two modules of this course are modified in a collaboration between the Research Division and another USAID operating unit (OU) to target that OU's desired audience.

MODULE 1: Research Evidence and the USAID Program Cycle (REAL-TIME VIRTUAL)
The goal of this module is to help participants understand the value of research evidence and learn how it can be used to inform decision-making throughout the USAID Program Cycle.

MODULE 2: Using Research Evidence at USAID (ONLINE)
As described above, this self-paced training takes approximately 1.5-2 hours to complete. In this blended course, it is a prerequisite for Module 3. The resource will also build participants' skills in finding, evaluating, and synthesizing research evidence from individual studies and broader bodies of research to inform decision-making.

8. CONCLUSION

This report is the first Capacity Assessment for Research, Evaluation, Statistics, and Other Analysis, required by The Foundations for Evidence-Based Policymaking Act (Pub. L. 115–435) "Evidence Act", to be submitted by Federal Agencies every four years, as part of their Strategic Plans. Data for the Capacity Assessment report was derived from (1) a review and analysis of USAID documents and tools for generating, managing, and using evidence, and (2) a capacity assessment study conducted by a third-party contractor, that assesses USAID's level of maturity to generate, manage and use evidence.

The report provided data on the (1) coverage, (2) quality, (3) methods, (4) effectiveness, and (5) independence of the statistics, evaluation, research, and analysis efforts of the Agency. In terms of coverage, USAID evaluation requirements allow for most USAID activities to be evaluated. In terms of research, in FY 2020, the Agency created and uploaded to its permanent repository a database of over 21,000 research and technical publications that were supported by USAID but that had not been previously deposited into the repository. In addition, USAID collects and publishes Development Assistance (DA) data and statistics on behalf of the U.S. Government and has many tools, including Foreign Aid Explorer, that are widely used as the main reference for U.S. development assistance data. To ensure that research evidence generated and used for making informed decisions is of high quality, USAID's Scientific Research Policy explicitly addresses quality standards in research.

The capacity assessment study assessed the level of maturity in the following areas, (1) resources, (2) culture, (3) collaborating, (4) learning, and (5) adapting - elements that are most critical to USAID evidence generation, management, and use. The study defined the maturity levels for elements assessed under each of these areas as (1) Basic, (2) Developing, (3) Advanced, and (4) Leading.

Under resources, USAID's maturity in capacity-building was assessed as developing. Leadership and advocacy were also assessed as developing as part of the culture of the Agency. In terms of collaborating, the assessment found that USAID was advanced in its internal collaboration but developing in its external collaboration. In terms of learning, the assessment ranks the availability of evidence to meet the needs of Operating Units, and the

existence of formal approaches for tracking evidence, as developing. In the area of adapting, the assessment raked USAID as advanced in both programmatic planning and adaptive management.

Based on these findings, the following next steps are being identified:

- Establishment of an annual peer-to-peer evidence exchange event that will bring together USAID staff, and Academics, think tanks, bilateral and multilateral donors, implementing partners (NGOs, contractors, etc.), evaluation and research firms, etc, to share and exchange with each other, the latest evidence and experience.

- Establish a quarterly report for USAID Missions to routinely provide updates on how evidence from evaluation, statistics, research, and other analysis are being applied in decision-making to make a difference in their programs.

- Explore the idea and feasibility of establishing a Global Development Evidence Bank, a public/private joint venture for a context driven search engine for multi-sectoral development evidence of what works, that can be accessed by all development practitioners. Internal discussions on this idea will be initiated in 2022.

- USAID is developing a knowledge management function in the PPL Bureau, through this will provide more guidance & resources for knowledge management and the development of simplified knowledge products, such as 1–2-page summaries with infographics and data visualizations.

- Explore the idea of developing training that is tailored to specific roles, such as leadership, supervisors, and managers.

These actions are expected to contribute towards addressing the gaps identified by respondents around engagement with external experts for evidence management and use. Bringing external partners together with USAID staff will initiate a process of relationship building that can lead to long-term benefits for evidence use and management. PPL/LER is currently working with internal stakeholders to develop approaches and a workplan for undertaking these next steps.

APPENDIX A

LIST OF ONGOING EVALUATIONS BY USAID OPERATING UNITS	
Barbados and Eastern Caribbean	Mid-term Evaluation of Youth Empowerment Services
Brazil	Mid Term /Late Term Evaluation of the Partnership to Conserve Amazon Biodiversity Program (PCAB) of IAA with USFS
Burma	Mid-term Performance Evaluation of Community Strengthening Project (CSP)
DDI - Center for Education	Impact Evaluation of READ Liberia
DDI - Center for Environment, Energy, and Infrastructure	Power Africa: Collaborating, Learning, and Adapting (CLA)
DDI - Gender Equality and Women's Empowerment Hub	Evaluation of the AMAA Activity in Malawi
DDI - Innovation, Technology, and Research Hub (ITR)	Randomized controlled trial of a family planning radio campaign in Burkina Faso
	Proactive Community Case Management and Child Survival - A Cluster RCT
	Liberia Cash Experiment
	DRC Cash Experiment
	ERIE Mixed Methods Retrospective Evaluation of the HESN Program
	ERIE Quasi-Experimental Evaluation of the PEER Program
	A cost-effective, scalable, self-sustainable model to deliver primary healthcare to urban poor
	The Cholera Hospital Based Intervention for 7 days (CHoBI 7) Trial
DDI - Local, Faith-based & Transformative Partnerships	Performance Evaluation of Local Works
Ethiopia	Final performance evaluation of the Value Chain Activity
Haiti	Final performance evaluation RANFOSE
Honduras	Honduras Local Governance Activity Performance Evaluation
Laos	Evaluation of USAID Nurture in Laos
	USAID Laos Lean to Read Evaluation
Liberia	Read Liberia EGRA Endline
Mali	PERFORMANCE EVALUATION OF Girls Leadership and Empowerment through Education (GLEE)
Nigeria	Agribusiness Investment Activity Mid-term evaluation
Pakistan	Pakistan Regional Economic Integration Activity - Project Performance Evaluation

Philippines	Final Performance Evaluation of Science, Technology, Research and Innovation for Development (STRIDE) Activity
Rwanda	Impact Evaluation comparing Household Grants Program and the Youth Employment program (Huguka Dukore Akazi Kanoze)
State Western Hemisphere Regional	ESC Impact Evaluation of Youth Empowerment Services
	Outreach Centers for Citizen Security Performance Evaluation
Tanzania	TUSHIRIKI PAMOJA -Promoting Accountable and Inclusive Politics Performance Evaluation
	Consortium for Elections and Political Process Strengthening (CEPPS)
USAID Asia Regional	Performance evaluation of the Building Healthy Cities activity
USAID Bureau for Conflict Prevention and Stabilization (CPS)	DCHA/CMM Conflict Assessment Task Order
USAID Bureau for Humanitarian Assistance (HA)	FFP - NRC EFSP Final Evaluation FY20
	FFP - ACTED EFSP Final Evaluation
	EMERGE Program Final Evaluation
USAID Bureau for Resilience and Food Security	Evaluation of the Partnership for Resilience and Economic Growth (PREG)
	WASH and nutrition - Cambodia
USAID East Africa Regional	Mid-Term Evaluation of the East African Community (EAC)
Zambia	End of Activity Performance Evaluation of the North Luangwa Ecosystem
Total	**38**