



**USAID**  
FROM THE AMERICAN PEOPLE

# META-EVALUATION AND SYNTHESIS REPORT

QUALITY REVIEW AND SYNTHESIS OF EVALUATIONS IN THE AFRICA  
REGION CONDUCTED BETWEEN 2017-2021

**DECEMBER 2022**

This publication was produced for review by the United States Agency for International Development. It was prepared by Elizabeth Mason with contributions from Anh Thu Hoang, Colin Maclean, Donna Podems, Ebenezer Owusu-Addo, and Nassrin Farzaneh.

# META-EVALUATION AND SYNTHESIS REPORT

## QUALITY REVIEW AND SYNTHESIS OF EVALUATIONS IN THE AFRICA REGION CONDUCTED BETWEEN 2017-2021

### PROJECT INFORMATION

**Project Title:** QUALITY REVIEW AND SYNTHESIS OF EVALUATIONS IN THE AFRICA REGION CONDUCTED BETWEEN 2017-2021

**Start and end dates:** August 2022 –December 2022

**Implementing Partners:** United States Department of Agriculture, Foreign Agriculture Services (USDA/FAS); University of Rhode Island (URI) and Kwame Nkrumah University of Science & Technology (KNUST), Ghana

**Coordinating Center:** Bureau of Integrated Rural Development (BIRD), College of Agriculture and Natural Resources, KNUST, Ghana

**Agreement number:** AFR-P-13-00001

**Funding:** USAID Africa Bureau

Team Leader	Elizabeth Mason
Senior Evaluation Specialists	Anh Thu Hoang Colin Maclean Donna Podems Ebenezer Owusu-Addo Nassrin Farzaneh
Points of Contact	Dr. Albert Arhin, BIRD/KNUST Alphonse Bigirimana, Africa Bureau Reid Hamel, Africa Bureau

# CONTENTS

<b>ACRONYMS AND ABBREVIATIONS</b> .....	<b>III</b>
<b>EXECUTIVE SUMMARY</b> .....	<b>IV</b>
INTRODUCTION .....	IV
METHODOLOGY .....	IV
OVERVIEW OF EVALUATIONS .....	V
QUALITY OF EVALUATIONS .....	VII
CONTENT OF EVALUATIONS .....	IX
<b>1. INTRODUCTION</b> .....	<b>I</b>
1.1. BACKGROUND .....	1
1.2. PURPOSE AND INTENDED USES .....	1
1.3. ASSESSMENT QUESTIONS .....	2
1.4. TEAM COMPOSITION .....	2
<b>2. METHODOLOGY</b> .....	<b>2</b>
2.1. DATA SOURCES .....	2
2.2. DATA COLLECTION .....	3
2.3. DATA ANALYSIS .....	4
2.4. REPORTING AND DISSEMINATION .....	5
2.5. LIMITATIONS AND RISKS .....	5
<b>3. OVERVIEW OF EVALUATIONS</b> .....	<b>6</b>
3.1. WHEN WERE EVALUATIONS PUBLISHED? .....	6
3.2. WHERE WERE EVALUATIONS COMMISSIONED? .....	7
3.3. WHAT WAS BEING EVALUATED? .....	9
3.4. HOW WERE EVALUATIONS CONDUCTED? .....	11
<b>4. QUALITY OF EVALUATIONS</b> .....	<b>17</b>
4.1. FINDINGS ON THE QUALITY OF EVALUATIONS .....	17
4.2. CONCLUSIONS ON THE QUALITY OF EVALUATIONS .....	26
<b>5. CONTENT OF EVALUATIONS</b> .....	<b>28</b>
5.1. COVERAGE OF THEMATIC TOPICS .....	28
5.2. FINDINGS ON KEY THEMATIC TOPICS .....	29
<b>ANNEX 1: STATEMENT OF WORK</b> .....	<b>45</b>
<b>ANNEX 2: EVALUATION REFERENCE LIST</b> .....	<b>50</b>
<b>ANNEX 3: LIST OF DOCUMENTS REFERENCED</b> .....	<b>60</b>
<b>ANNEX 4: DATA COLLECTION TOOLS</b> .....	<b>61</b>
EVALUATION QUALITY CHECKLIST .....	61
EVALUATION QUALITY CHECKLIST RATERS' GUIDE .....	65
EVALUATION DESCRIPTIVE DATA AND CONTENT TOOL .....	72
EVALUATION DESCRIPTIVE DATA AND CONTENT TOOL - RATER'S GUIDE .....	77

## FIGURES

Figure 1: Evaluations by year.....	6
Figure 2: Timing of evaluations.....	6
Figure 3: Evaluations commissioned by region.....	7
Figure 4: Map of evaluation commissioners.....	8
Figure 5: Scope of the evaluations.....	9
Figure 6: Sector coverage in evaluations.....	9
Figure 7: Sector coverage in evaluations of integrated activities (n=15).....	10
Figure 8: Overlap of integrated activities.....	10
Figure 9: Evaluation type.....	11
Figure 10: Number of key evaluation questions.....	11
Figure 11: Evaluation questions.....	12
Figure 12: Performance evaluations with cause-and-effect questions and methodologies.....	12
Figure 13: Data collection methods.....	14
Figure 14: Data collection approaches.....	14
Figure 15: Distribution of in-person and virtual data collection.....	15
Figure 16: Virtual data collection over time by method.....	15
Figure 17: Data analysis methods.....	16
Figure 18: Distribution of evaluation quality scores.....	17
Figure 19: Distribution of evaluation quality scores by year.....	18
Figure 20: Average E3 sector evaluation quality score by year, 2009 - 2021.....	18
Figure 21: Percent of quality factors in each quality tier.....	19
Figure 22: Sustainability in evaluation reports.....	29
Figure 23: Coverage of ToC lessons in evaluations.....	38
Figure 24: Principles for effective local capacity development and strong government engagement.....	41

## TABLES

Table 1: Assessment Methodology.....	iv
Table 2: 10-point quality score factors.....	3
Table 3: Evaluation commissioners.....	8
Table 4: Evaluation budget metrics.....	13
Table 5: Evaluation quality rankings for all factors.....	20
Table 6: Evaluation quality factor rankings.....	26
Table 7: Coverage of thematic topics in evaluations.....	28
Table 8: Examples of evaluations addressing sustainability.....	30
Table 9: Levels of government engagement in evaluations.....	40

## ACRONYMS AND ABBREVIATIONS

ADS	Automated Directives System
AFR	USAID/Africa Bureau
Ag	Agriculture
BIRD	Bureau of Integrated Rural Development
CAE	Cause-and-effect
COVID-19	Coronavirus disease of 2019
DRG	Democracy, Human Rights, and Governance
E3	USAID/Bureau for Economic Growth, Environment, and Education
E&I	Energy and Infrastructure
Ed	Education
EQ	Evaluation Question
FDG	Focus Group Discussion
GE	Government Engagement
KNUST	Kwame Nkrumah University of Science and Technology
LCD	Local Capacity Development
LOP	Life of the Project
M&E	Monitoring and Evaluation
MSI	Management Systems International
NRM	Natural resource management
PPL	USAID/Policy, Planning, and Learning
SOW	Statement/Scope of Work
T&I	Trade and Investment
ToC	Theory of Change
OU	Operating Unit
US	United States
USAID	US Agency for International Development
WASH	Water, Sanitation, and Hygiene

# EXECUTIVE SUMMARY

## Introduction

This assessment was commissioned by the USAID Africa Bureau (AFR) to contribute to the knowledge base on the quality and content of evaluations conducted in sub-Saharan Africa from 2017 to 2021. It includes both a meta-evaluation component that examines the quality of evaluation reports and a meta-analysis, or synthesis, component that consolidates findings across evaluations on key topics.

The assessment aims to raise awareness among Africa Mission and Africa Bureau about their evaluation quality to inform strategies for improvement, which may include developing guidance, products, technical support, or conducting additional analysis. The assessment also aims to provide key findings on project design, implementation, and performance in the Africa Region with the intent to inform ongoing and future programming through this report.

This independent assessment has been prepared by the Bureau of Integrated Rural Development of the Kwame Nkrumah University of Science and Technology (KNUST), under a partnership between the United States Department of Agriculture/Foreign Agricultural Service, University of Rhode Island, and KNUST. The technical team consisted of a Team Leader and five Senior Evaluation Specialists. All work was done remotely.

## Methodology

This assessment analyzed the 177 evaluations funded or commissioned by USAID Operating Units in the Africa Region completed between 2017 and 2021. The assessment team also referred to previous studies and secondary documents for background information and tool development, and to inform analysis.

Table I below summarizes the assessment questions, data collection methods, and data analysis methods. The team underwent an extensive training and calibration process to ensure consistent application of the data collection tools and minimize Interrater Reliability issues.

*Table I: Assessment Methodology*

Assessment Question	Data Collection	Data Analysis
1. What do Africa evaluations completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?	<b>Evaluation Quality Checklist</b> , a tool developed in 2012 for a USAID study of evaluation quality and subsequently used for other meta-evaluations.	<ul style="list-style-type: none"><li>• Descriptive statistics of quality factors</li><li>• Calculation of composite score, and analysis compared to prior studies</li></ul>
2. What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?	<b>Descriptive Data and Content Extraction Tool</b> , based on prior USAID evaluation synthesis studies and adapted to address key themes of interest	<ul style="list-style-type: none"><li>• Descriptive statistics of quantitative questions</li><li>• Synthesis of qualitative data for trends or key points</li></ul>

This assessment included four deliverables, designed to contribute the highest possible understanding, dissemination, and utilization of its results.

Assessment Report	Briefing Note	Spreadsheet	Webinar
Narrative of the assessment team’s findings and conclusions	Summary of the assessment team’s findings and conclusions	Tool for further exploration of the dataset and learning for stakeholders	Presentation for AFR Bureau, AFR/Missions, and other key stakeholders

The limitations of this assessment include a limited timeframe and the content of the evaluations themselves.

This assessment was conducted on a short timeframe due to contractual constraints, which required the assessment team to prioritize either data extraction or report writing. To address this limitation, the assessment was designed to extract as much information as possible from the evaluations and a spreadsheet deliverable was added. This report focuses on quantitative findings (Sections 3 and 4), with a brief presentation on key qualitative topics (Section 5).

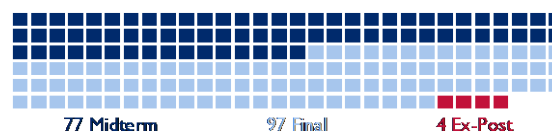
The assessment team also recognized the limitation of content extraction on key themes due to the content of the evaluations themselves. The topics chosen for this synthesis by AFR are of particular interest to the Bureau to inform current and future programming. However, an evaluation synthesis can only present findings based on the evaluations reviewed and these topics may have greater or lesser coverage in any given evaluation. The assessment team acknowledged this limitation by being clear on the sources and depth of evidence.

The primary risks to this assessment were the potential for variation of ratings between team members and potential conflicts of interest. To mitigate the risk of discrepancies in the application of the assessment tools, the team underwent a robust training and calibration process. To mitigate the risk of any potential conflict of interest between ratings and personal involvement in evaluations, team members did not score evaluations they worked on.

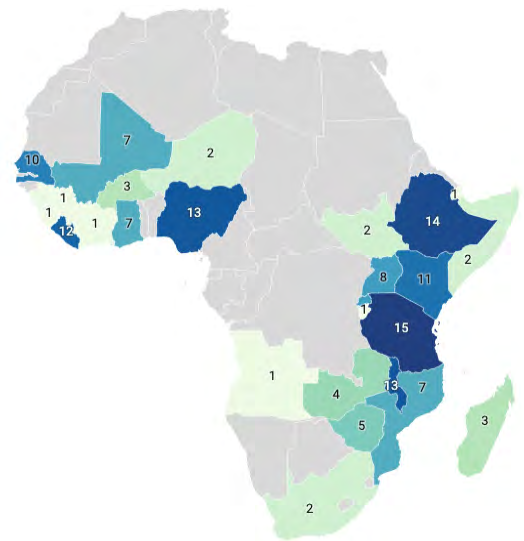
## Overview of Evaluations

The average number of evaluations was 35 per year, with a range from 43 to 49 per year in 2017-2019, and 25 and 16 evaluations in 2020 and 2021 respectively during the COVID-19 pandemic. The average budget for performance evaluations was \$266k, and the average for impact evaluations was \$1.3m.

Evaluations were conducted at three points in time – midterm evaluations during the intervention, final evaluations towards the end of the activity, and ex-post evaluations at some point after the activity had ended. This assessment included 77 midterm (44 percent), 97 final (55 percent), and 4 ex-post (2 percent) evaluations.

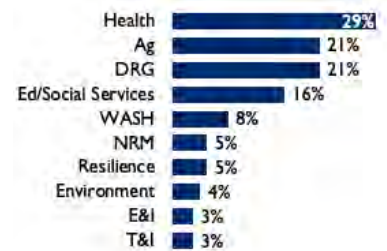


The highest number/volume of evaluations was commissioned in the West Africa region (70 evaluations, 40 percent) and East Africa region (67 evaluations, 38 percent), followed by the Southern Africa region (39 evaluations, 22 percent). This assessment also included one cross-region evaluation of a Power Africa intervention commissioned out of Washington. The vast majority of evaluations (154 evaluations, 87 percent) were commissioned by country Missions. Twenty-two (12 percent) were commissioned by regional Missions. The number of evaluations ranged from one to 15 per country/regional mission.

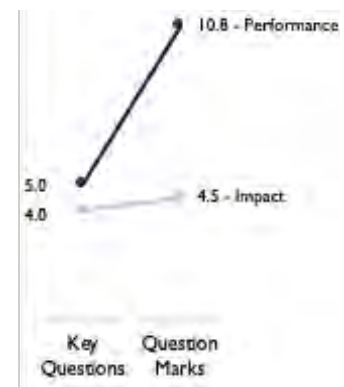


The vast majority of evaluations (157 evaluations, 89 percent) examined a single activity. Of these, 94 percent were of an activity in a single country and 6 percent were in multiple countries. To a lesser extent, evaluations also examined projects or programs (11 evaluations, 6 percent) and multiple activities (8 evaluations, 5 percent). One evaluation examined the intervention of a global initiative in a single country.

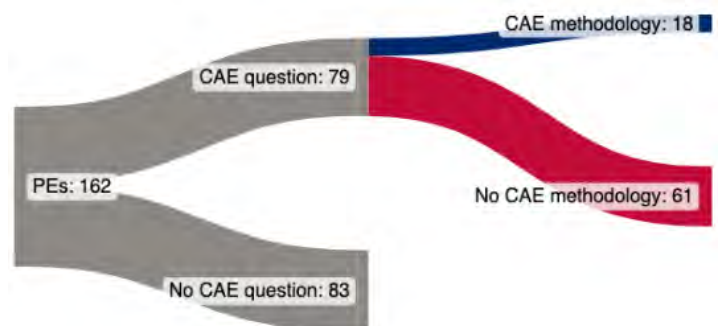
The largest share of evaluations was commissioned by the health sector (29 percent); followed by agriculture and food security (21 percent) and Democracy, Human Rights, and Governance (DRG, 21 percent). Fifteen evaluations (eight percent) evaluated an integrated activity, which span more than one sector and were funded by multiple USAID technical units. The figure to the right sums to more than 100 percent, as integrated activities are commissioned by more than one sector. Fourteen percent of evaluations addressed an activity that addressed either youth skill building and/or youth civic engagement.



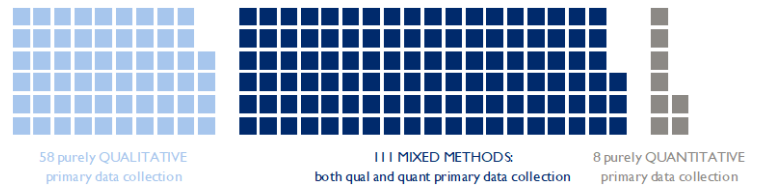
The vast majority of evaluations in the assessment were performance evaluations (162 evaluations, 92 percent). Fifteen (eight percent) were impact evaluations. Across all evaluations, the average count of key questions was five. The count of all question marks was 10. This average is primarily driven by performance evaluations as shown in the line graph to the right. The difference between key and total question marks for impact evaluations is minimal (4.0 vs. 4.5), whereas performance evaluations included more than twice as many question marks as key questions, on average.



Of the 162 performance evaluations, almost half (79 evaluations, 49 percent) were asked to address a cause-and-effect question. However, only 23 percent of those (18 evaluations) included a cause-and-effect methodology to try to answer the question such as pre-post designs or contribution analysis. The remaining 77 percent (61 evaluations) did not include a cause-and-effect methodology. This distribution is shown in the chart to the right (CAE – cause-and-effect).



Sixty-three percent of evaluations used a mixed methods approach, using both qualitative and quantitative data collection methods. Thirty-three percent of evaluations used only qualitative primary data collection methods, and eight used only quantitative primary methods.



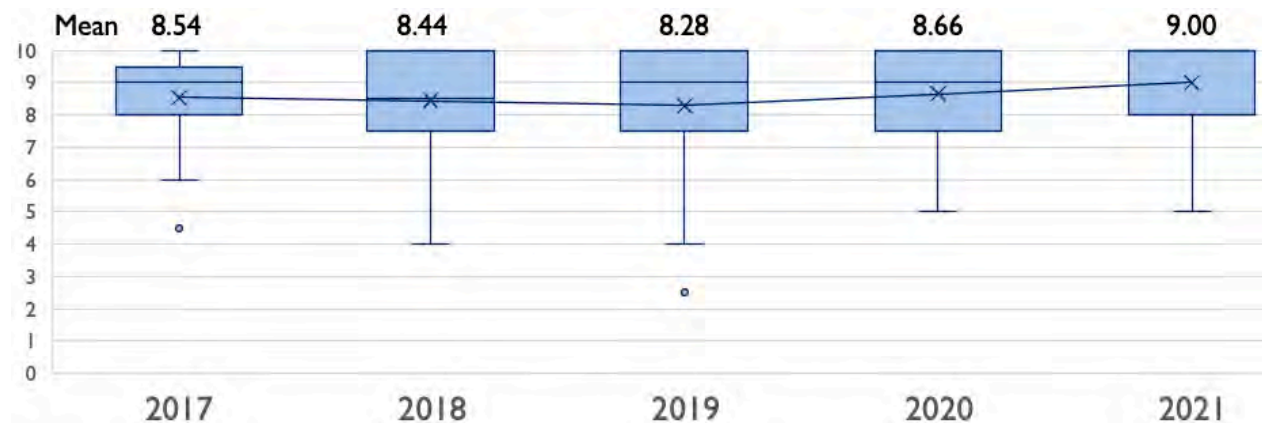
## Quality of Evaluations

*Question 1: What do Africa evaluations (evaluations funded/commissioned by USAID Operating Units in the Africa Region) completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?*

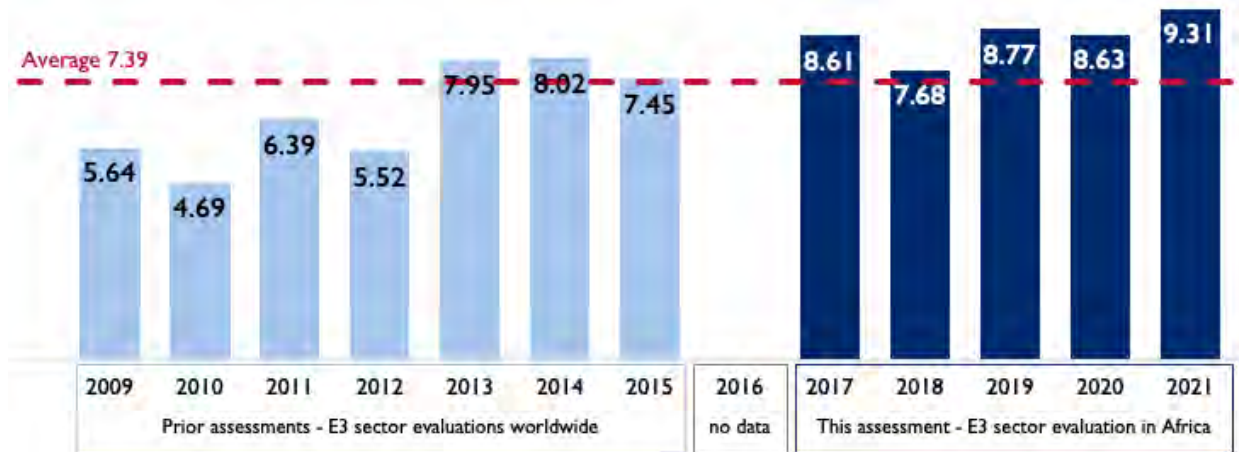
To assess the quality of the evaluation reports included in this study, the assessment team applied the USAID Evaluation Quality Checklist. Eleven of the 45 items on the checklist are combined to create a composite quality score.

### Findings

The mean (average) quality score of the 2017-2021 Africa evaluations was 8.51, with a range from of 8.28 in 2019 to 9.00 in 2021. The box and whisker plot shows the distribution of quality scores by year. The “x” is the mean score by year, with the line connecting each x showing the trend over time. The depth of the boxes shows that there was little variation in quality in 2017, but that, in general, quality scores are converging on the upper end of the 10-point scale between 2018 and 2021.



This follows the improvement trend seen in the Bureau for Economic Growth, Environment, and Education (E3) Evaluation Syntheses assessments, which included scores for E3 sector evaluations from the earlier Bureau for Policy, Planning, and Learning (PPL) study (2009-2012) and the E3 assessment years (2013-2015) for evaluations worldwide. While these scores are not directly comparable because the samples are different (worldwide vs. Africa), the general trend of improvement in quality scores that was observed in the E3 assessments holds true in this assessment.



Consistent with the PPL assessment, the assessment team grouped these factors into four quality tiers by the percent of evaluations rated positively for each factor:



Forty seven percent (27 factors) were rated as good, 38 percent (17 factors) as fair, 2 percent (1 factor) as marginal. While 13 percent (6 factors) were rated as weak, not all of these would necessarily need to be addressed in an evaluation, such as including a statement of difference (factor number 37). Additionally, some factors may no longer be relevant, such as including a description of how data will be transferred to USAID (factor number 38).



## Conclusions

The overall quality of Africa evaluations has improved according to the Evaluation Quality Checklist. The evaluation quality scores increased as compared to previous studies and are clustered on the top end of the 10-point quality scale. Areas of improvement can be grouped into structural/clarity issues and methodology issues. While the responsibility for most of these factors sits primarily with the evaluation team, they are also affected by how USAID writes SOWs, manages the evaluation process, and reviews draft evaluation reports. The assessment team concludes that these areas are where USAID could focus efforts to improve evaluation quality.

Easier areas for improvement for evaluation teams may be those factors related to the structure or basic clarity of the evaluation report, such as ensuring clarity between the SOW and report questions,

including COI forms, improving team descriptions defining data collection and analysis methods by question, clarifying recommendations, including sources and data collection instruments as annexes, and describing the activity’s theory of change at the beginning of the report. Additional areas for structural improvement by USAID include incorporating the required language from the ADS mandatory reference *Criteria to Ensure the Quality of the Evaluation Report* in evaluation of SOWs.

Other areas for improvement may be more complicated to achieve, such as those where evaluation methodology or the implementation of the methodology needs to be addressed by the evaluation team. These include increasing the inclusion of se-disaggregated data and addressing gender differential access and results; including an evaluation specialist and local team members on the evaluation team; distinguishing clearly between findings, conclusions, and recommendations throughout the report; making a clear connection between findings, conclusions, and recommendations; and writing transparently so that readers can follow a clear connection between findings and their sources.

The assessment team’s final conclusion on evaluation quality is on the Evaluation Quality Checklist tool itself. The tool was designed in 2012 and was structured around USAID’s 2011 Evaluation Policy. Evaluations have shown great improvements in the past decade to achieving the principles outlined in the policy and its successors. This trend has been documented by this assessment, as well as those in the E3 evaluation synthesis studies. The tool remains a valid one to track improvement over time and to highlight opportunities for improvement, such as those listed above. But with scores converging at the top end of the 10-point scale, it may be time to broaden the definition of quality and for future meta-evaluations to delve into a more detailed analysis to advance learning and push the quality of evaluations to a higher level.

## Content of Evaluations

*Question 2: What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?*

To answer this question, the assessment team developed the descriptive data and content tool to extract programmatic lessons. Five key thematic topics were analyzed by the Senior Evaluation Specialists: sustainability; monitoring systems, learning, and adaptive management; government engagement; beneficiary selection and targeting; and theory of change. Their findings and conclusions are presented in full in the body of the report. The other topics were not analyzed for this report due to time constraints of the contract.

The coverage of thematic topics ranged from 85 percent (sustainability) to 12 percent (youth topics). COVID-19 was addressed in 12 percent of all evaluations. However, one would not expect to see references to COVID-19 in the 2017-2019 evaluations before the pandemic began. COVID-19 was addressed in 51 percent of evaluations published in 2020 and 2021.

Percent of evaluations addressing key topics	
Sustainability	85%
Monitoring systems, learning, and adaptive management	72%
Government engagement	66%
Beneficiary selection and targeting	56%
Theory of Change	48%
Performance target setting	47%
Innovation	23%
Unanticipated results	15%
Youth capacity building	12%
Youth civic engagement and leadership	12%
COVID-19	12%

## **Sustainability**

The issue of sustainability is complex and often clearly misunderstood regarding how it should be considered in terms of how projects or beneficiaries relate to it. In most cases, as there is no way of knowing whether continuity of outcomes prevails post-project unless there is follow-up by USAID, most evaluations essentially offered recommendations that they believed should be undertaken by the project or other parties to ensure sustainability. Furthermore, in both mid-term and final evaluations, there was little distinction drawn between interventions and outcomes though they are inextricably linked. In terms of project design, though more a recommendation than a lesson learned, projects would be better served if sustainability plans were embedded in Theories of Change or equivalent.

## **Monitoring, Learning, and Adaptive Management**

The majority (71 percent) of the evaluations reviewed discussed some aspect of the monitoring system, learning, and adaptive management components of the USAID activities/projects in the Africa region - although with varying degrees of focus on each sub-component (i.e., monitoring system or learning or adaptive management). Of these, only 6 percent specifically addressed adaptive management issues. Whilst the analysis depicts instances of good practices, evaluations revealed the need for greater diligence in optimizing the design, implementation, and performance of the monitoring system, learning, and adaptive management components of the interventions.

## **Theory of Change (ToC)**

Three conclusions can be drawn from the findings relating to the ToCs. First, relatively few evaluations pay attention to the program ToCs, which are themselves supposed to inform the evaluation design. Second, ToC potential is seen as supporting programs' innovation and adaptation in response to dynamic contexts. It is thus important to make ToCs flexible and contextually responsive both in the design stage and, more importantly, in implementation and performance management. Finally, as assumptions take time to become more visible, an iterative, staged ToC process that is embedded within program learning can inspire innovations and improvements in program strategies and strengthen the potential of programs to optimize their development outcomes.

## **Government Engagement (GE)**

The meta-analysis found that activities with high government engagement seemed to have strong collaborative relationships with the IP and/or other stakeholders. In these evaluations, there were reports of: trust between government, project and/or other actors; local (government) appreciation for the Activity; and project ownership. These evaluations also reported regular meetings and information sharing among actors throughout the LOP, including “co-managing” field activities with government actors. In general projects with weak government engagement did not start with the local system, neither did they align with local priorities, nor appreciate local capacities. Indeed, these evaluations reported poor communication/interactions between the IP and government stakeholders which likely prohibited mutual learning and collaboration-strengthening.

## **Beneficiary Targeting and Selection**

The meta-analysis revealed a lack of data in evaluation related to the topic of beneficiary selection and targeting which suggests that either (1) beneficiary selection is not an identified area of focus for USAID evaluations and/or (2) monitoring or other data suggest few issues or challenges with beneficiary selection or targeting. A key lesson is that evaluations should address how and why beneficiary groups were selected and how that selection relates to the problem statement and the theory of change if this is an area of interest.

# I. INTRODUCTION

## I.1. Background

Evaluation utilization can be supported by both meta-analysis and meta-evaluation, both of which have strong histories within the US Agency for International Development (USAID). This assessment was commissioned by the USAID Africa Bureau (AFR) to contribute to the knowledge base on the quality and content of evaluations conducted in sub-Saharan Africa from 2017 to 2021. It includes both a meta-evaluation component that examines the quality of evaluation reports (Section 4) and a meta-analysis, or synthesis, component that consolidates findings across evaluations on key topics (Section 5).

USAID has conducted several meta-evaluations to assess the quality of its evaluation reports at the global, bureau, and sectoral levels. The first study of evaluation quality based on the 2011 Evaluation Policy was the *Meta-Evaluation of the Quality and Coverage of Evaluation Reports*,<sup>1</sup> commissioned by USAID/Policy, Planning and Learning (PPL). This study assessed the quality and coverage of 340 randomly selected evaluations completed between 2009 and 2012. The Evaluation Rating Checklist developed for this study has subsequently been used by other USAID evaluation quality assessments to track progress over time, including the 2015<sup>2</sup> and 2017<sup>3</sup> *E3 Sectoral Evaluation Synthesis* reports.

USAID also values learning from evaluations through meta-analysis, or evaluation synthesis. As described in the *PPL Discussion Note: Making Evidence Accessible through Evaluation Synthesis Platforms*,<sup>4</sup> “Evaluation syntheses aim to consolidate what is being learned from evaluations and make that evidence more readily available. These studies can help busy decision-makers – who may not have time to locate and review all relevant evaluations and research reports – use existing evidence to better address pressing development challenges, consistent with USAID’s commitment to supporting countries.”

## I.2. Purpose and Intended Uses

AFR expects that the results of this assessment will raise awareness among Africa Mission and Africa Bureau leadership about the strengths and weaknesses of the evaluations produced across the Bureau. Findings on factors affecting evaluation quality will help the Bureau and Missions to devise strategies for evaluation quality improvement which may include developing additional guidance, products, and technical support.

AFR hopes that the review will spark interest among Missions to analyze and address internal factors that affect the quality of evaluations. During the dissemination of findings from this assessment, Missions will be encouraged to analyze other factors not covered by the assessment that influence the quality of

### Evidence Syntheses for Decision-Makers

**Meta-analyses** extract and present the substantive findings of evaluations. Meta-analyses can also consolidate evidence from other types of studies.

**Meta-evaluations** evaluate evaluations for their compliance with Agency policy, or their quality and evidence strength.

Source: *Learning from Evaluation Syntheses*

<sup>1</sup> <https://www.usaid.gov/evaluation/meta-evaluation-quality-and-coverage>

<sup>2</sup> [https://www.usaid.gov/sites/default/files/documents/1865/E3\\_Sectoral\\_Synthesis\\_Report.pdf](https://www.usaid.gov/sites/default/files/documents/1865/E3_Sectoral_Synthesis_Report.pdf)

<sup>3</sup> [https://pdf.usaid.gov/pdf\\_docs/PA00MP17.pdf](https://pdf.usaid.gov/pdf_docs/PA00MP17.pdf)

<sup>4</sup> Page 1, [https://usaidearninglab.org/sites/default/files/resource/files/dn-evidence\\_eval\\_synthesis\\_final2021\\_1.pdf](https://usaidearninglab.org/sites/default/files/resource/files/dn-evidence_eval_synthesis_final2021_1.pdf)

evaluations, including internal Mission operations (evaluation planning, staff capacity, evaluation independence, evaluation management, etc.).

In addition, this assessment aims to provide key findings across evaluations on project design, implementation, and performance in the Africa Region. Both the Bureau and Missions will gain easier access to key substantive lessons learned in diverse technical sectors with the intent to inform ongoing and future programming through this report.

### **1.3. Assessment Questions**

This assessment sought to answer the questions below, as provided in the assignment statement of work (SOW, Annex 1: Statement of Work). These questions were developed by AFR to be consistent with the questions of previous evaluation quality and synthesis assessments.

1. What do Africa evaluations (evaluations funded/commissioned by USAID Operating Units in the Africa Region) completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?
2. What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?

### **1.4. Team Composition**

This independent assessment has been prepared by the Bureau of Integrated Rural Development of the Kwame Nkrumah University of Science and Technology (KNUST), under a partnership between the United States Department of Agriculture/Foreign Agricultural Service, University of Rhode Island, and KNUST. The technical team consisted of a Team Leader and five Senior Evaluation Specialists. All the work was done remotely.

## **2. METHODOLOGY**

### **2.1. Data Sources**

This assessment focused on evaluations completed between 2017 and 2021 that were funded or commissioned by USAID Operating Units (OUs) in the Africa Region. An initial review by AFR of the Evaluation Registry compiled a list of 232 evaluations in this period. Of these, the assessment team identified 177 evaluation reports. The remaining reports listed on the Evaluation Registry were not evaluations (i.e., assessments, special studies), were not within the assessment time period (i.e., listed in the Evaluation Registry as 2017, but report indicated 2016), or could not be located (i.e., no link in the evaluation registry, a link to incorrect report, or a broken link with no recovery success via an internet search). These evaluations and their evaluation numbers used throughout this report are included as Annex 2: Evaluation Reference List.

The assessment team also referred to previous studies and other secondary documents for background information and tool development, and to inform analysis. These documents are cited in Annex 3: List of Documents Referenced.

## 2.2. Data Collection

### 2.2.1. Tool Development

The assessment team developed two data collection tools to answer its assessment questions: an evaluation quality checklist and a descriptive/content extraction tool. The tools and their raters' guides are included in this report as Annex 4: Data Collection Tools.

#### Evaluation Quality Checklist Tool

To answer Question I about evaluation quality, this assessment adopted the Evaluation Rating Checklist. This tool, developed by Management Systems International (MSI) under a PPL/LER contract for the *Meta-Evaluation of the Quality and Coverage of Evaluation Reports*, is based on USAID evaluation quality standards and practices described in the 2011 Evaluation Policy and ADS 201. It was subsequently used by MSI to conduct a Sectoral Evaluation Synthesis of E3 Bureau evaluations completed between 2013-2014 (first study)<sup>5</sup> and fiscal year 2015 (second study)<sup>6</sup>. The checklist has undergone several tests, including inter-rater reliability tests. The raters' guide also helps the team to score evaluations consistently.

The tool includes 42 factors related to USAID evaluation quality factors/criteria, of which 11 factors are used to calculate an overall score of the quality of each evaluation.<sup>7</sup> Shown in Table 2, the score includes a full point for nine factors, and a half point each for two factors, for a possible score from 0 to 10.

Table 2: 10-point quality score factors

Category	Factors from the Evaluation Quality Checklist
Executive Summary	1. Is there an Executive Summary which accurately reflects the most critical elements of the report?
Program/Project Background	2. Are the basic characteristics of the program, project or activity described (title, dates, funding organization, budget, implementing organization, location/map, target group)? [half point]
	3. Is the program or project's "theory of change" described (intended results (in particular the project purpose); development hypotheses; assumptions) [half point]
Methodology	8. Does the report (or methods annex) describe specific data collection methods the team used?
	10. Does the report (or methods annex) describe specific data analysis methods the team used?
Study Limitations	16. Does the report include a description of study limitations (lack of baseline data; selection bias as to sites, interviewees, comparison groups; seasonal unavailability of key informants)?
Findings	20. Did the findings presented appear to be drawn from social science data collection and analysis methods the team described in its study methodology (including secondary data it assembled or reanalyzed)?
	23. Are findings clearly distinguished from conclusions and recommendations in the report, at least by the use of language that signals transitions ("the evaluation found that...", "the team concluded that...")?
Recommendations	32. Are all the recommendations supported by the findings and conclusions presented? (Can a reader follow a transparent path from findings to conclusions to recommendations?)
Annexes	33. Is the evaluation SOW included as an annex to the evaluation report?
	35. Are data collection instruments provided as evaluation report annexes?

<sup>5</sup> [https://www.usaid.gov/sites/default/files/documents/1865/E3\\_Sectoral\\_Synthesis\\_Report.pdf](https://www.usaid.gov/sites/default/files/documents/1865/E3_Sectoral_Synthesis_Report.pdf)

<sup>6</sup> [http://pdf.usaid.gov/pdf\\_docs/PA00MPI7.pdf](http://pdf.usaid.gov/pdf_docs/PA00MPI7.pdf)

<sup>7</sup> For more information on the Evaluation Rating Checklist and 10-point score, see Part 2, Annex B of *Meta-Evaluation of the Quality and Coverage of Evaluation Reports*

## **Descriptive Data and Content Tool**

To answer Question 2 about the content of evaluations, the assessment team developed a descriptive and content data collection tool and raters' guide. This tool was adapted from tools developed by MSI for the E3 evaluation syntheses to capture data on key themes, lessons learned, and evaluation quality factors not included in the quality checklist. The assessment team adapted the E3 tool to focus on AFR's priority programming themes and refined it throughout the tool training process, described below.

### **2.2.2. Tool Training and Team Calibration**

To ensure the team had a consistent understanding of the tools and the scoring process, the assessment Team Leader trained the team in the application of the tools. All team members reviewed an initial set of evaluations and individual scoring was compared to identify any major variations in scores between team members, and to clarify the team's understanding of elements of the tool with inconsistent scores. After this first calibration round, the Team Leader revised the descriptive data and content tool to improve clarity between team members and to better capture topics of interest to AFR.

The assessment team then undertook two more rounds of calibration, during which team members reviewed a second and third set of evaluations in pairs. Groups met to discuss and agree on scoring and the Team Leader reviewed the scoring for any differences. At this stage, the team members were at 97 percent agreement on data points related to the evaluation quality score (43 out of 44 data points). Additional guidance was provided by the team leader on any points of divergence.

### **2.2.3. Scoring and Data Extraction**

During tool administration, the Team Leader assigned each evaluation to one team member. Team members read and scored the evaluations over a six-week period. Data was entered into an online data collection form, creating a master spreadsheet of all data collected. The Team Leader periodically checked in with team members to assess progress, ensure quality control, and answer any questions. Team members were also encouraged to reach out to each other throughout scoring and data extraction to maintain calibration.

## **2.3. Data Analysis**

Following completion of tool administration, the Team Leader cleaned the data set and followed up with team members for any clarifications needed on their ratings.

The Team Leader analyzed the quantitative data. Analysis included running descriptive statistics of data from the Evaluation Quality Checklist Tool and the quantitative measures from the Descriptive Data and Content Tool for the entire sample, as well as descriptive statistics disaggregated by time, location, sector, and evaluation methodology.

For analysis of the qualitative data captured through the Descriptive Data and Content Tool, AFR selected five topics that were of the most interest. The five Senior Evaluation Specialists each analyzed one topic, assessing report content and synthesizing trends or key points of learning. These topics are explored in detail under Section 5.2. Findings on Key Topics.

## 2.4. Reporting and Dissemination

This assessment included four deliverables, designed to contribute the highest possible understanding, dissemination, and utilization of its results.

Assessment Report	Briefing Note	Spreadsheet	Webinar
Narrative of the assessment team's findings and conclusions	Summary of the assessment team's findings and conclusions	Tool for further exploration and learning for stakeholders	Presentation for AFR Bureau, AFR/Missions, and other key stakeholders

## 2.5. Limitations and Risks

### 2.5.1. Limitations

The limitations of this assessment include a limited timeframe and the content of the evaluations themselves.

This assessment was conducted on a short timeframe due to contractual constraints, which required the assessment team to prioritize either data extraction or report writing. To address this limitation, the assessment was designed to extract as much information as possible from the evaluations and a spreadsheet deliverable was added to support future analysis of the data collected. This report focuses on quantitative findings (Sections 3 and 4), with a brief presentation on key qualitative topics (Section 5).

The assessment team also recognized the limitation of content extraction on key themes due to the content of the evaluations themselves. The topics chosen for this synthesis by AFR are of particular interest to the Bureau to inform current and future programming. However, an evaluation synthesis can only present findings based on the evaluations reviewed and these topics may have greater or lesser coverage in any given evaluation. The assessment team acknowledged this limitation by being clear on the sources and depth of evidence.

### 2.5.2. Risks

The primary risks to this assessment were the potential for variation of ratings between team members and potential conflicts of interest.

To mitigate the risk of discrepancies in the application of the assessment tools, the team underwent a robust training and calibration process as described above in Section 2.2.2.

To mitigate the risk of any potential conflict of interest between ratings and personal involvement in evaluations, team members did not score evaluation that they worked on.

### 3. OVERVIEW OF EVALUATIONS

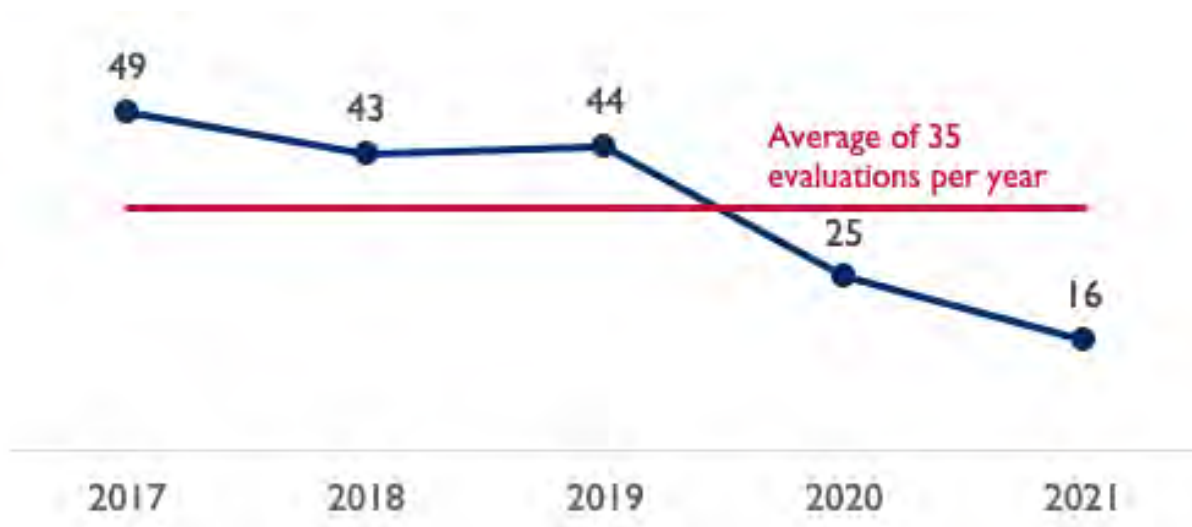
The following section describes the evaluations included in the assessment, detailing when they were published, where they were commissioned, what was being evaluated, and how they were conducted.

#### 3.1. When were evaluations published?

##### 3.1.1. Evaluations by year

The average number of evaluations was 35 per year. However, the evaluation counts from before compared to during the COVID-19 pandemic are distinctive. The number of evaluations ranged from 43 to 49 per year in 2017-2019. In 2020 and 2021, 25 and 16 evaluations were conducted.

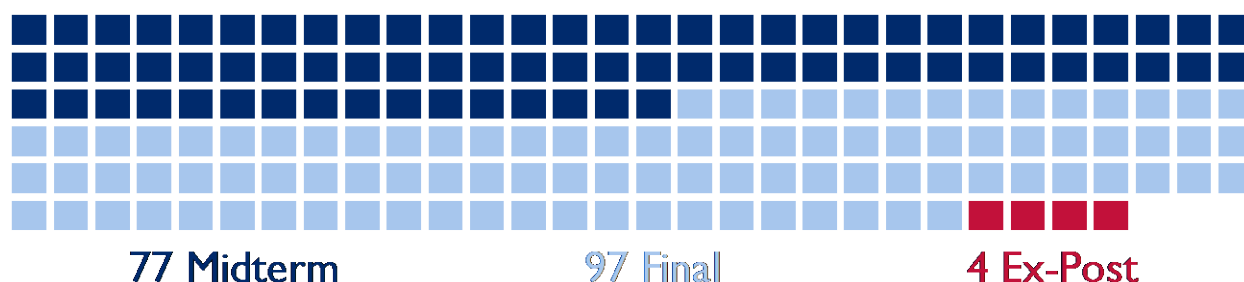
Figure 1: Evaluations by year



##### 3.1.2. Timing of Evaluations

Evaluations were conducted at three points in time – midterm evaluations during the intervention, final evaluations towards the end of the activity, and ex-post evaluations at some point after the activity had ended. This assessment included 77 midterm (44 percent), 97 final (55 percent), and 4 ex-post (2 percent) evaluations.

Figure 2: Timing of evaluations

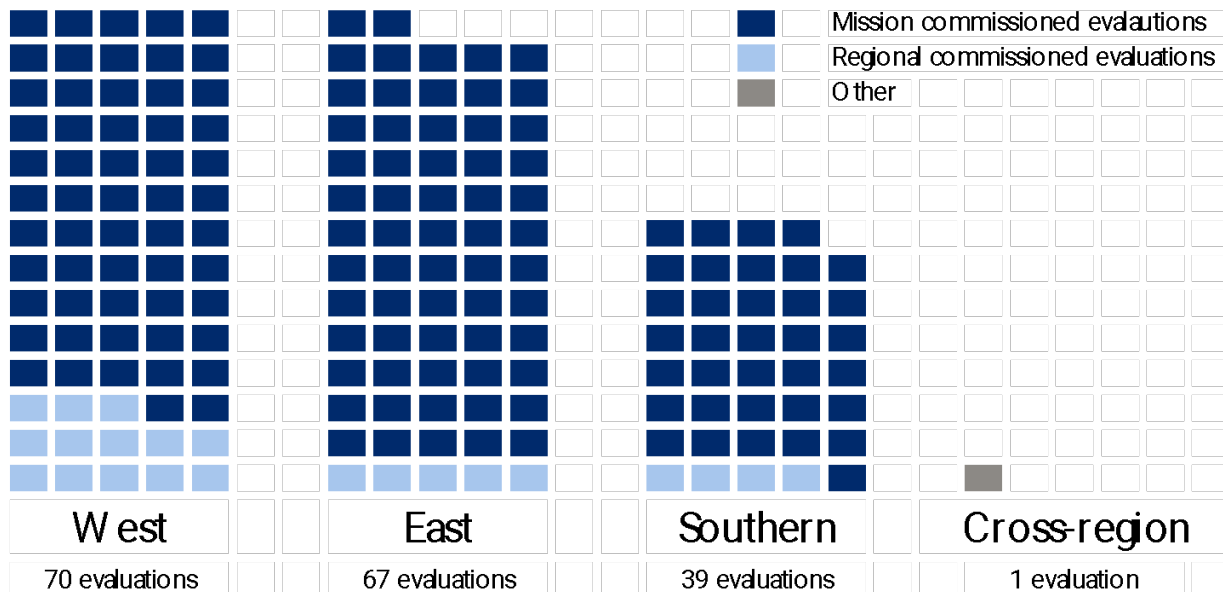


## 3.2. Where were evaluations commissioned?

### 3.2.1. Region

The highest number/volume of evaluations was commissioned in the West Africa region (70 evaluations, 40 percent) and East Africa region (67 evaluations, 38 percent), followed by the Southern Africa region (39 evaluations, 22 percent). This assessment also included one cross-region evaluation of a Power Africa intervention commissioned out of Washington.<sup>8</sup> The vast majority of evaluations (154 evaluations, 87 percent) were commissioned by country Missions, represented in dark blue in the figure below. Twenty-two (12 percent) were commissioned by regional Missions.

Figure 3: Evaluations commissioned by region



<sup>8</sup> Evaluation I6

### 3.2.2. Mission

As shown in Table 3 and Figure 4, the number of evaluations ranged from one to 15 per country/regional mission. The highest number of evaluations were commissioned in Tanzania (15), Ethiopia (14), Nigeria (13), Malawi (13), the West Africa Regional Mission (12), Kenya (11), Liberia (12), and Senegal (10).

Figure 4: Map of evaluation commissioners

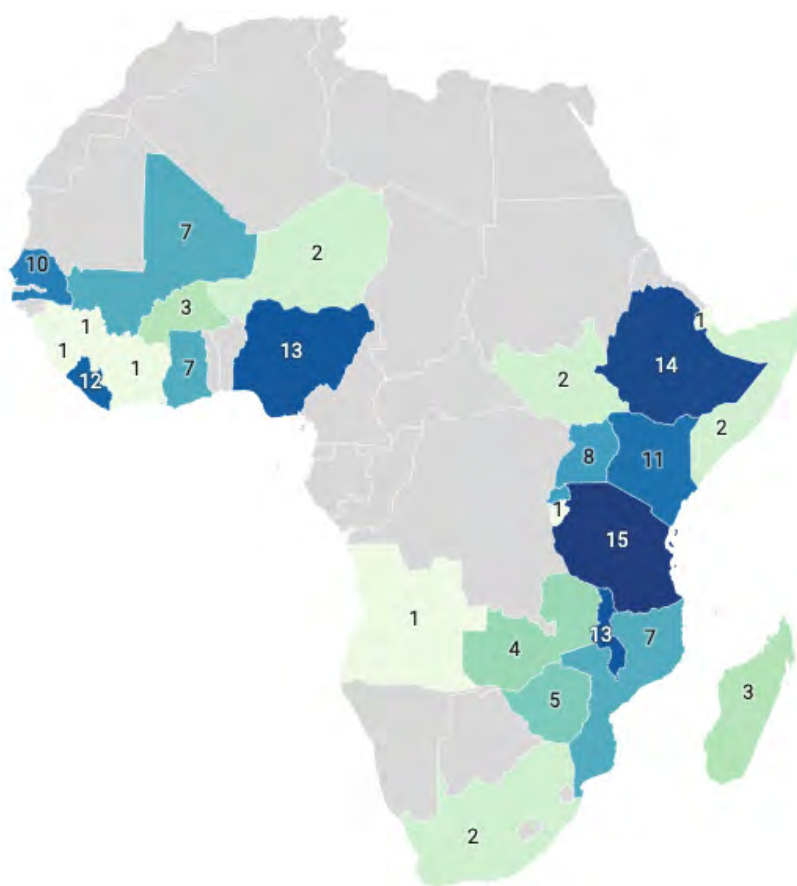


Table 3: Evaluation commissioners

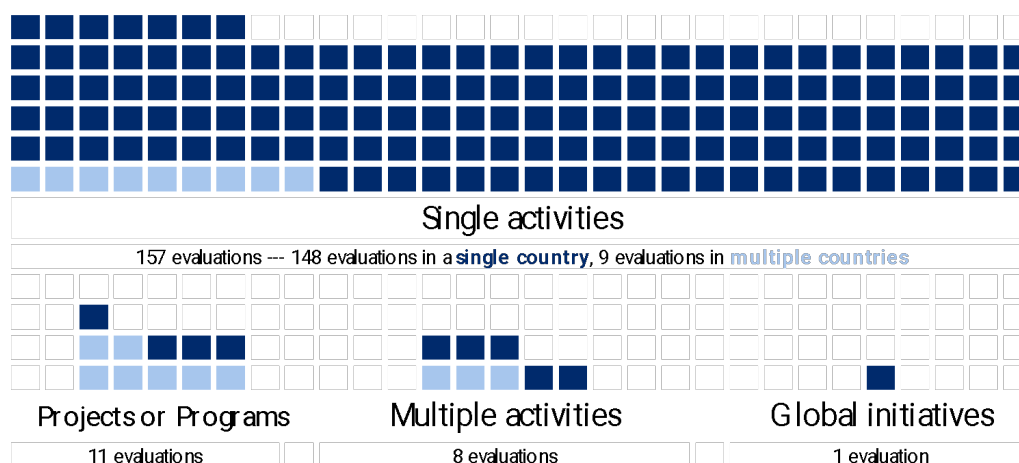
Commissioner	Count	Percent
<b>West</b>	<b>70</b>	<b>40%</b>
Nigeria	13	7%
Liberia	12	7%
WA Regional	12	7%
Senegal	10	6%
Mali	7	4%
Ghana	7	4%
Burkina Faso	3	2%
Niger	2	1%
Sierra Leone	1	1%
Cote d'Ivoire	1	1%
Sahel Regional	1	1%
Guinea	1	1%
<b>East</b>	<b>67</b>	<b>38%</b>
Tanzania	15	8%
Ethiopia	14	8%
Kenya	11	6%
Uganda	8	5%
Rwanda	8	5%
EA Regional	4	2%
South Sudan	2	1%
Somalia	2	1%
Djibouti	1	1%
CA Regional	1	1%
Burundi	1	1%
<b>Southern</b>	<b>39</b>	<b>22%</b>
Malawi	13	7%
Mozambique	7	4%
Zimbabwe	5	3%
Zambia	4	2%
SA Regional	4	2%
Madagascar	3	2%
South Africa	2	1%
Angola	1	1%

### 3.3. What was being evaluated?

#### 3.3.1. Evaluations by scope

The vast majority of evaluations (157 evaluations, 89 percent) examined a single activity. Of these, 94 percent were of an activity in a single country and 6 percent were in multiple countries. To a lesser extent, evaluations also examined projects or programs (11 evaluations, 6 percent) and multiple activities (8 evaluations, 5 percent). One evaluation examined the intervention of a global initiative in a single country.

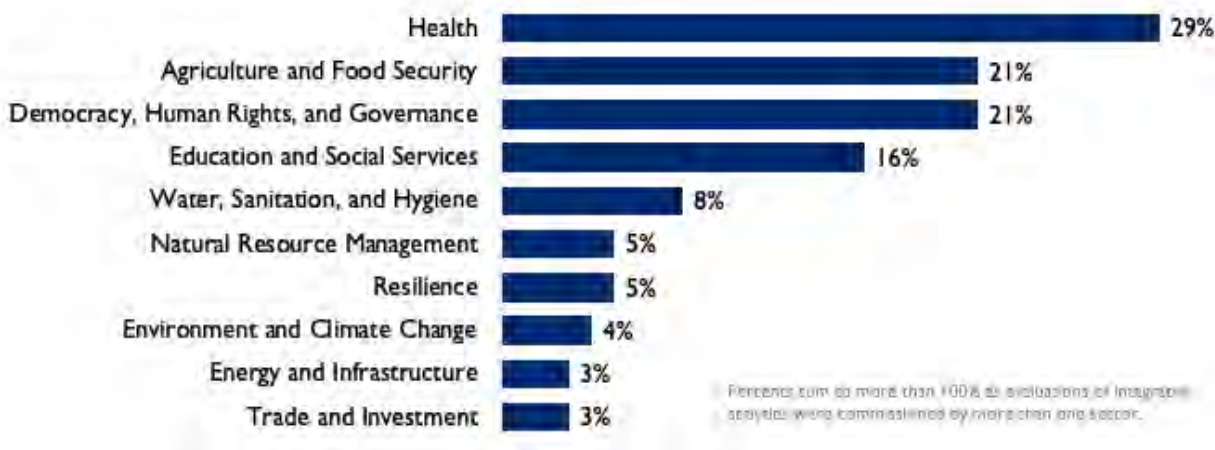
Figure 5: Scope of the evaluations



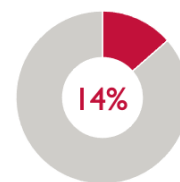
#### 3.3.2. Evaluations by sector

The largest share of evaluations was commissioned by the health sector (29 percent), followed by agriculture and food security (21 percent) and Democracy, Human Rights, and Governance (DRG, 21 percent). Coverage for all sectors is shown in the figure below. The figure sums to more than 100 percent, as integrated activities (discussed in the next section) are commissioned by more than one sector.

Figure 6: Sector coverage in evaluations

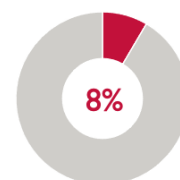


This assessment was specifically interested in extracting information on two youth-specific intervention types: youth skill building and youth civic engagement, which were addressed in 24 evaluations (14 percent). Twenty-one evaluations (12 percent) were of activities that included youth skill building and/or other capacity strengthening support with the aim of increasing youth employment. Of these, seven evaluations (33 percent) discussed a positive association between the activity's interventions and youth employment. Twenty-one evaluations (12 percent) were of activities that addressed youth civic engagement and/or leadership.



### 3.3.3. Evaluations of integrated activities

Fifteen evaluations (eight percent) were of an integrated activity. Integrated, or multi-sectoral activities, span more than one sector and were funded by multiple USAID technical units. Twelve activities spanned two sectors, two activities spanned three sectors, and one activity spanned four sectors.



As shown in Figure 7, around half of the integrated activities addressed health (53 percent) and/or agriculture and food security (47 percent). About a third addressed water, sanitation, and hygiene (WASH, 33 percent) and resilience (27 percent). One fifth addressed DRG and environment and climate change (20 percent each). Less common sectors for integrated activities included education and social services (13 percent); natural resource management (7 percent); and trade and investment (7 percent).

Figure 7: Sector coverage in evaluations of integrated activities (n=15)

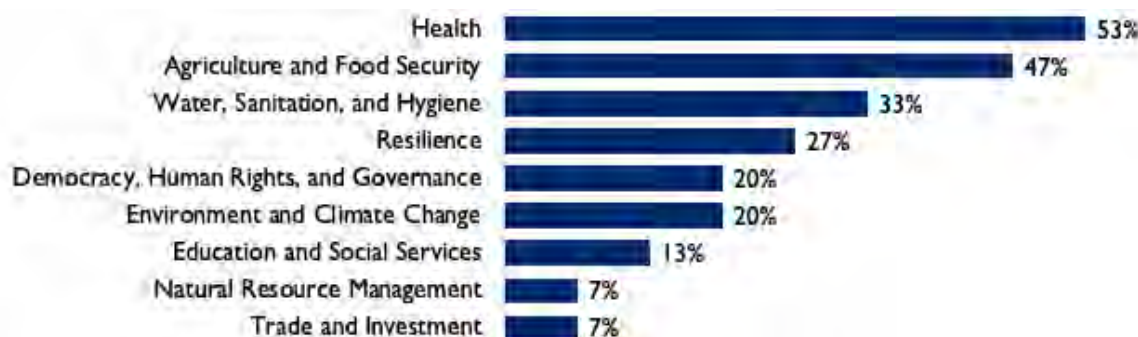


Figure 8 shows the two-way overlap between integrated activities. For example, of the eight integrated activities in the health sector, five were integrated with agriculture, three WASH, one resilience, two DRG, and one education/social services. The highest level of overlap was between the health and the agriculture and food security sectors.

Figure 8: Overlap of integrated activities

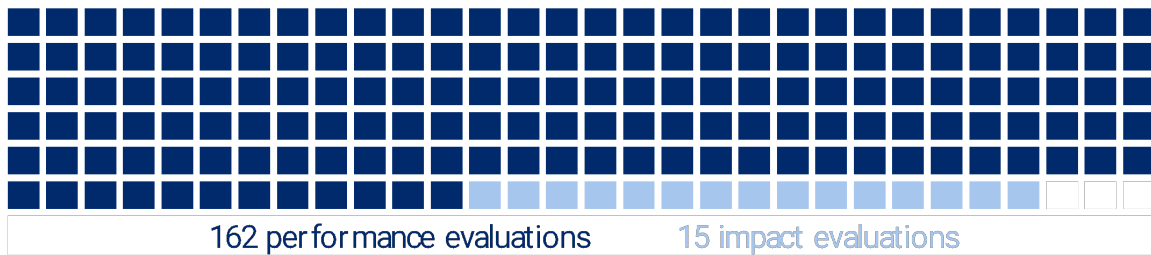
	Health	Ag/ Food Security	WASH	Resilience	DRG	Envirom/ Climate	Ed/ Social Services	NRM	Trade
Health	8 evals	5	3	1	2	-	1	-	-
Ag/ Food Security	5	7 evals	3	-	1	-	-	-	1
WASH	3	3	5 evals	1	2	-	-	-	-
Resilience	1	-	1	4 evals	-	2	-	-	-
DRG	2	1	2	-	3 evals	-	1	-	-
Envirom/ Climate	-	-	-	2	-	3 evals	-	1	-
Ed/ Social Services	1	-	-	-	1	-	2 evals	-	-
NRM	-	-	-	-	-	1	-	1 eval	-
Trade	-	1	-	-	-	-	-	-	1 eval

### 3.4. How were evaluations conducted?

#### 3.4.1. Evaluation approach – performance vs. impact

The vast majority of evaluations in the assessment were performance evaluations (162 evaluations, 92 percent). Fifteen (eight percent) were impact evaluations.

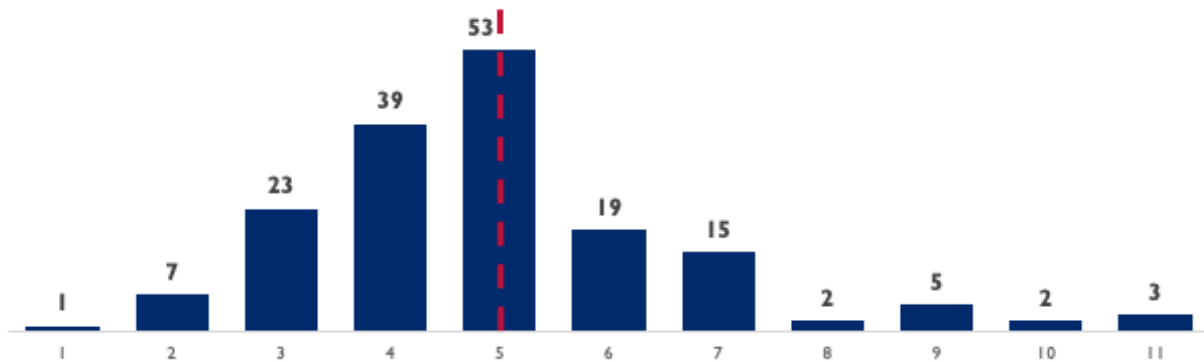
Figure 9: Evaluation type



#### 3.4.2. Evaluation questions

Ninety-five percent of evaluations (169) addressed questions. The remaining five percent (eight evaluations) dealt with issues or topics. According to the *PPL How-To Note on Preparing Evaluation Statements of Work*,<sup>9</sup> evaluations should strive to address three to five “key questions.” as shown in Figure 10, the key evaluation question average in this sample was five and its distribution was approximately normal.

Figure 10: Number of key evaluation questions



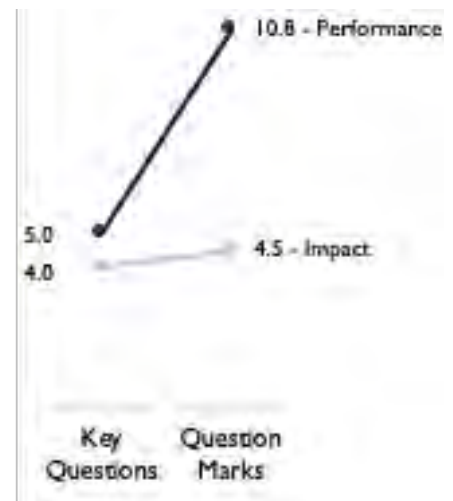
<sup>9</sup> [https://www.usaid.gov/sites/default/files/documents/1865/508\\_Howto\\_Note\\_Evaluation\\_Statements\\_Work.pdf](https://www.usaid.gov/sites/default/files/documents/1865/508_Howto_Note_Evaluation_Statements_Work.pdf)

This assessment counted both the number of key questions and of question marks, which capture sub-questions or when a key question contains more than one question. For example, this assessment addressed **two key questions** and **three question marks**:

1. What do Africa evaluations completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?
2. What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?

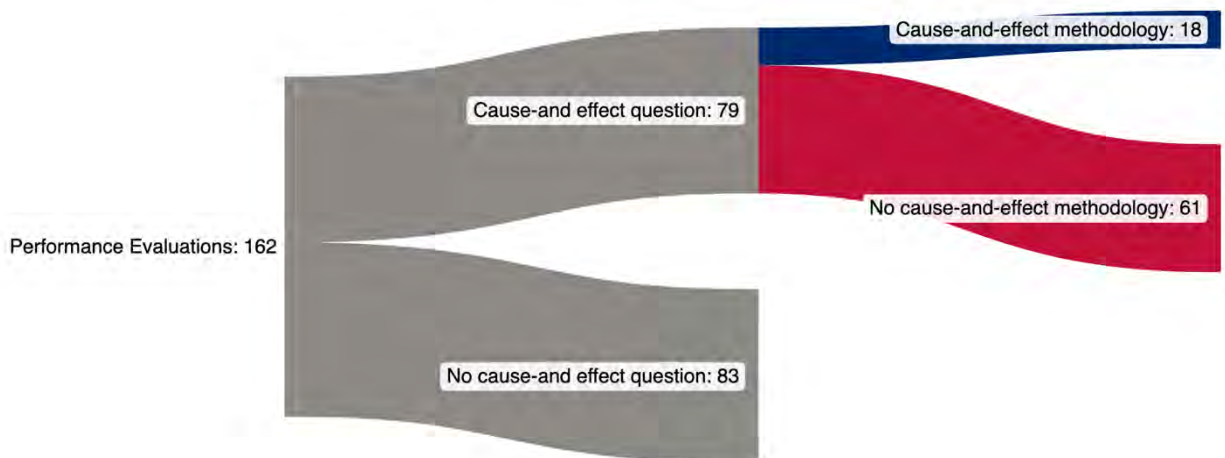
Across all evaluations, the average count of key questions was five. The count of all question marks was 10. This average is primarily driven by performance evaluations as shown in Figure 11. The difference between key and total question marks for impact evaluations is minimal (4.0 vs. 4.5), whereas performance evaluations included more than twice as many question marks as key questions, on average.

Figure 11: Evaluation questions



Of the 162 performance evaluations, almost half (79 evaluations, 49 percent) were asked to address a cause-and-effect question.<sup>10</sup> However, only 23 percent of those (18 evaluations) included a cause-and-effect methodology to try to answer the question such as pre-post designs or contribution analysis. The remaining 77 percent (61 evaluations) did not include a cause-and-effect methodology, shown in red in Figure 12.

Figure 12: Performance evaluations with cause-and-effect questions and methodologies



<sup>10</sup> Evaluation questions are included in the spreadsheet deliverable to enable future analysis.

### 3.4.3. Evaluation budgets

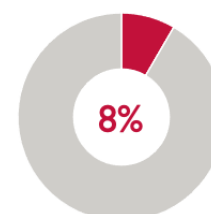
Across the 157 evaluations with available budget data,<sup>11</sup> the average budget was \$354,000. Total budgets ranged from \$11,000 to \$3.6m. As shown in Table 4, impact evaluations had larger budgets than performance evaluations for all metrics (mean, median, lowest budget, and highest budget).

Table 4: Evaluation budget metrics

Budget Metric	Performance Evaluations	Impact Evaluations	Percent difference
Number of evaluations	144	13	n/a
Highest	\$ 1,500,000	\$ 3,628,179	142%
Median	\$ 265,537	\$ 1,328,650	400%
Mean	\$ 250,000	\$ 885,772	254%
Lowest	\$ 11,500	\$ 300,000	2,509%

### 3.4.4. Participatory evaluations

Fourteen evaluations (eight percent) stated that they included a participatory approach. Of those, two stated that beneficiaries participated in the design process, three in data collection, and five during the formulation of recommendations. Five evaluations stated that beneficiaries were involved in another stage of the evaluation or stated that beneficiaries participated in the evaluation but did not indicate how.



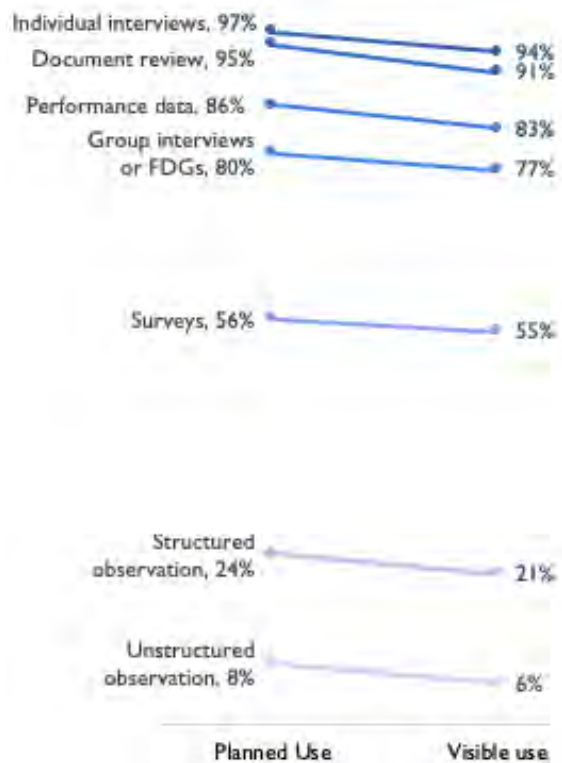
<sup>11</sup> Evaluation budgets were sourced from the Evaluation Registry, which included data for 160 of the 177 evaluations. Three evaluations were excluded from the budget analysis for having higher than probable budget figures (\$55M, \$18M, and \$5M). The assessment team did not independently verify the budget numbers.

### 3.4.5. Data collection methods

The most common data collection method with demonstrated use in the findings of evaluations reports was individual interviews (94 percent), followed by document review (91 percent), collection of activity performance data (83 percent), group interviews or focus group discussions (FDGs, 77 percent), and surveys (55 percent). The least used methods of data collection were observations, with 21 percent of evaluation using structured observation and 6 percent using unstructured observation.

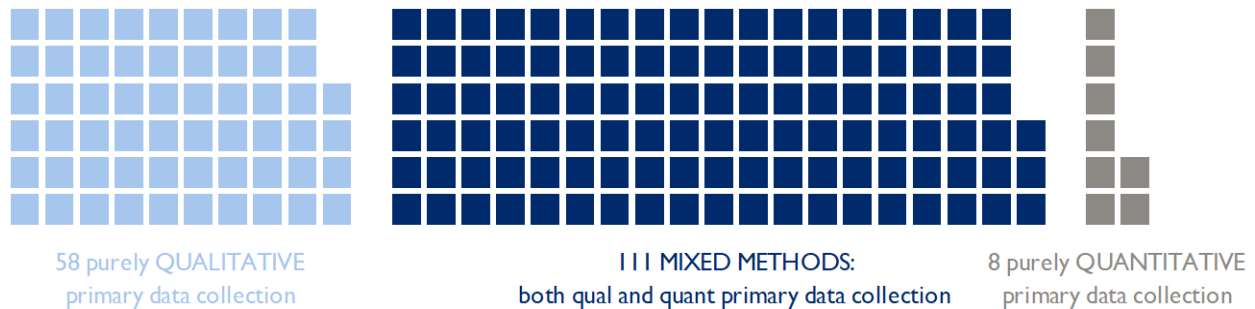
As shown in Figure 13, there was a decrease of two to four percentage points between the stated intent to use a specific method (shown on the left in the figure) and the demonstrated use of that method in the findings (shown on the right). This decrease indicates that either the data were not collected, or that the evaluation team did not clearly present the methodological underpinnings of their findings.

Figure 13: Data collection methods



The assessment team also examined whether evaluations planned to use quantitative primary data collection methods (surveys and structured observation), qualitative primary data collection methods (individual and group interviews, FDGs, and unstructured observations), or a combination of both. As shown in Figure 14: Data collection approaches, 63 percent of evaluations used a mixed methods approach, using both qualitative and quantitative data collection methods. Thirty-three percent of evaluations used only qualitative primary data collection methods, and eight used only quantitative primary methods.

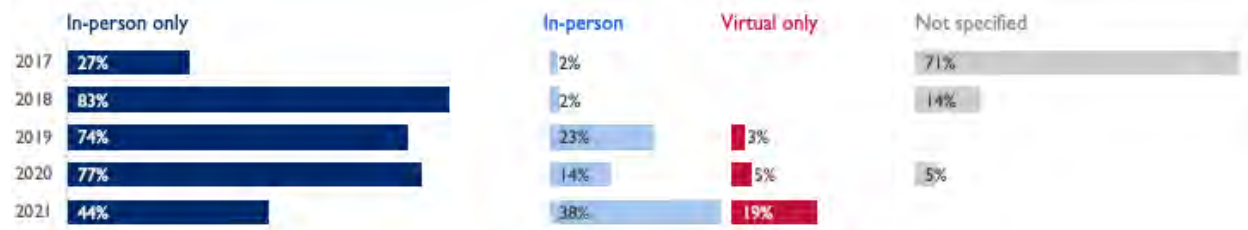
Figure 14: Data collection approaches



The assessment team looked at how data were collected to determine whether there was a shift from in-person to virtual during the COVID-19 pandemic. The team found that the proportion of virtual data collection increased substantially in 2021 relative to prior years. Figure 15 details these findings on whether data collection was **in-person**, **virtual**, or a **mix of both** for those evaluations that used these three data collection methods. It also details the proportion of evaluations that used the data collection method but did **not specify** whether it was in-person or virtual.

Figure 15: Distribution of in-person and virtual data collection

**Individual Interviews**



**Group Interviews and FGDs**

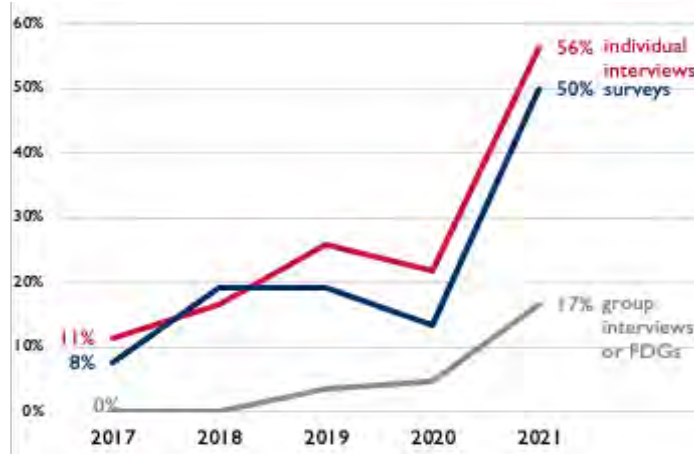


**Surveys**



Analyzing just the evaluations that clearly stated how data was collected, the proportion of interviews conducted virtually increased from 11 to 56 percent between 2017 and 2021 as shown in Figure 16. Virtual administration of surveys increased from 8 to 50 percent. In 2017 and 2018, no evaluations cited conducting FGDs or group interviews virtually while in 2021, 17 percent of those using this data collection method did.

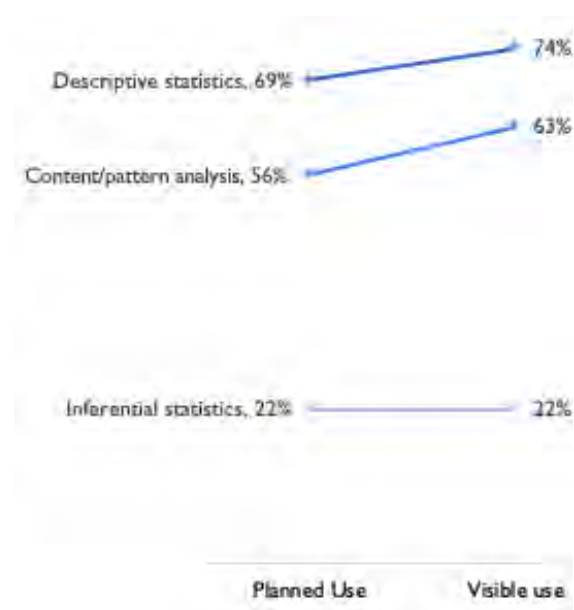
Figure 16: Virtual data collection over time by method



### 3.4.6. Data analysis methods

The most common data analysis method with demonstrated use in evaluations was descriptive statistics (74 percent), followed by content or pattern analysis (63 percent), and inferential statistics (22 percent). In contrast with data collection methods, the frequency of data analysis methods that were demonstrably used in the body of the reports was the same as, or higher than, the methods described in the methodology section of the reports, as shown in Figure 17. This increase in visible use over planned use for descriptive statistics and content or pattern analysis indicates that the reader can identify how the data were analyzed in the findings of the report. However, the approach to the analysis was not written clearly enough in neither the methodology section (in the body of the report) nor the methodology annex for the reader to understand.

Figure 17: Data analysis methods



## 4. QUALITY OF EVALUATIONS

*Question 1: What do Africa evaluations (evaluations funded/commissioned by USAID Operating Units in the Africa Region) completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?*

To assess the quality of the evaluation reports included in this study, the assessment team applied the USAID Evaluation Quality Checklist. Eleven of the 45 items on the checklist are combined to create a composite quality score. See Section 2.2.1 and Table 2 for more details on the checklist tool.

### 4.1. Findings on the Quality of Evaluations

#### 4.1.1 Average Evaluation Quality Score

The mean (average) quality score of the 2017-2021 Africa evaluations was 8.51. Evaluation quality scores clustered on the higher end of the 10-point scale with a median of 9.00. As shown in Figure 18 below, the distribution of quality scores is left-skewed – the number of evaluations increased as scores increased.

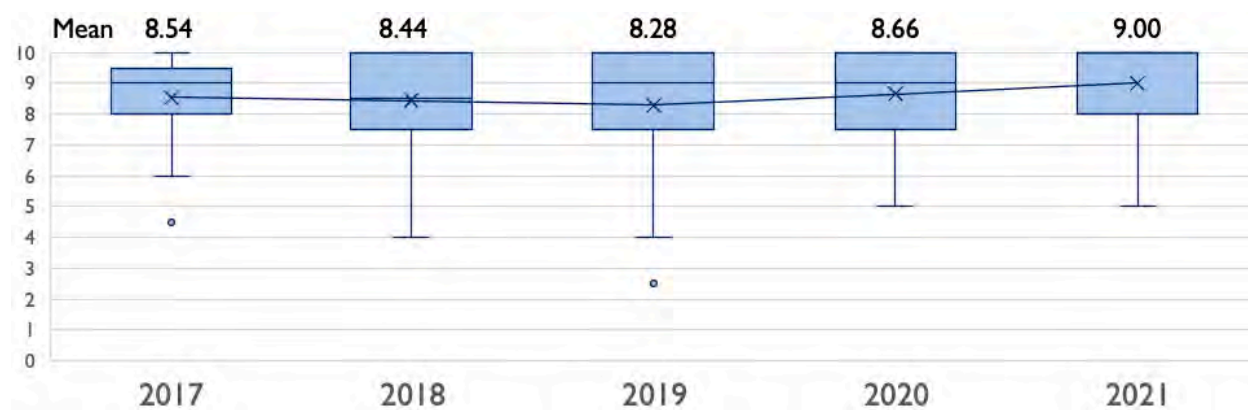
Figure 18: Distribution of evaluation quality scores



### 4.1.2 Change in Evaluation Quality Score over Time

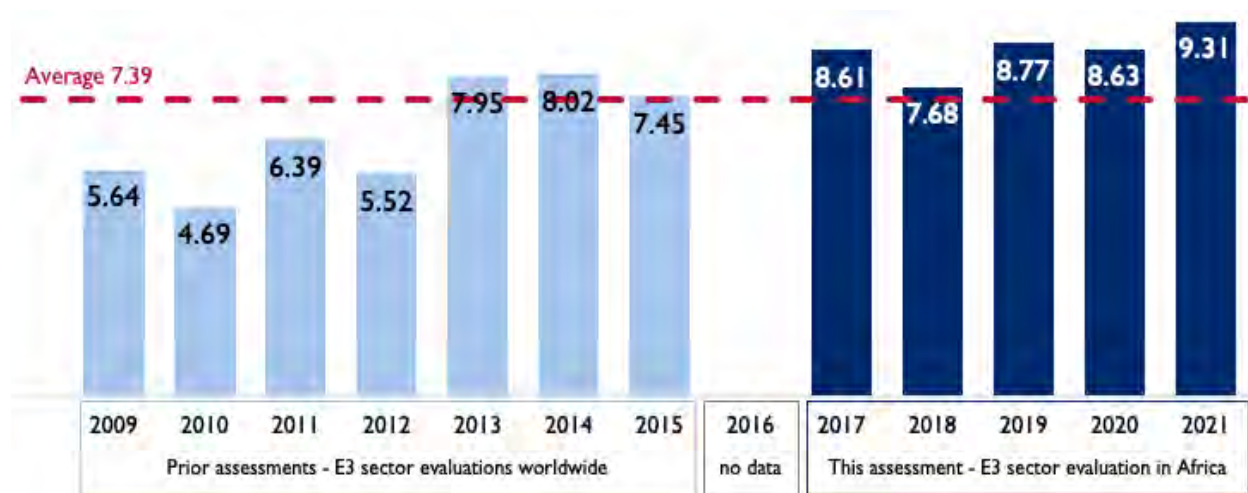
Average quality scores ranged slightly from a low of 8.28 in 2019 to a high of 9.00 in 2021. The following box and whisker plot shows the distribution of quality scores by year. The “x” is the mean score by year, with the line connecting each x showing the trend over time. The depth of the boxes shows that there was little variation in quality in 2017, but that, in general, quality scores are converging on the upper end of the 10-point scale between 2018 and 2021.

Figure 19: Distribution of evaluation quality scores by year



To assess how the quality of AFR evaluations has changed since prior studies, the assessment team compared the quality scores to the E3 evaluation syntheses. The E3 assessments included scores for E3 sector evaluations from the earlier PPL study (2009-2012) and the E3 assessment years (2013-2015) for evaluations worldwide. These scores are shown in light blue in the figure below. The scores of this assessment’s evaluations (2017-2021) are shown in dark blue for E3 sectors only.<sup>12</sup> While these scores are not directly comparable because the samples are different (worldwide vs. Africa), the general trend of improvement in quality scores that was observed in the E3 assessments holds true in this assessment.

Figure 20: Average E3 sector evaluation quality score by year, 2009 - 2021



<sup>12</sup> E3 related sectors with evaluations in this assessment were agriculture and food security; education and social services; energy and infrastructure; environment and climate change; natural resource management; and trade and investment.

### 4.1.3 Evaluation quality: strengths and weaknesses

This section analyzes the individual factors on the Evaluation Quality Checklist. Consistent with the PPL assessment, the assessment team grouped these factors into four quality tiers by the percent of evaluations rated positively for each factor:

<b>Weak</b> Less than 25 percent of evaluations scored positively	<b>Marginal</b> 25 – 49 percent of evaluations scored positively	<b>Fair</b> 50 – 79 percent of evaluations scored positively	<b>Good</b> 80 percent or higher of evaluations scored positively
--	---	---	--

Table 5 and Figure 21 summarizes the scoring of all quality factors, which are discussed in detail in the following sections. Forty seven percent (27 factors) were rated as good, 38 percent (17 factors) as fair, 2 percent (1 factor) as marginal. While 13 percent (6 factors) were rated as weak, not all of these would necessarily need to be addressed in an evaluation, such as including a statement of difference (factor number 37). Additionally, some factors may no longer be relevant, such as including a description of how data will be transferred to USAID (factor number 38).

Figure 21: Percent of quality factors in each quality tier



Table 5: Evaluation quality rankings for all factors

Item from the Evaluation Quality Checklist	Rated positively
<b>Executive Summary</b>	
1. Executive summary mirrors critical report elements	86%
<b>Activity/Program/Project Background</b>	
2. Project characteristics described	92%
3. Project "theory of change" described	77%
<b>Evaluation Purpose</b>	
4. Management purpose described	93%
<b>Evaluation Questions</b>	
5. Questions were linked to purpose	99%
6. Questions in report same as in SOW	68%
7. Written approval for question changes documented	14%
<b>Methodology</b>	
8. Data collection methods described	100%
9. Data collection methods linked to questions	62%
-- . Gender sensitive data collection methods	33%
10. Data analysis methods described	80%
11. Data analysis methods linked to questions	60%
<b>Team Composition</b>	
12. External team leader	65%
13. Evaluation team included an evaluation specialist	55%
14. Evaluation team included local members	58%
15. Conflict of interest forms signed	50%
<b>Study Limitations</b>	
16. Study limitations included	95%
<b>Responsiveness to Evaluation Questions</b>	
17. Report structured to respond to questions	88%
18. Evaluation questions addressed in body of report	91%
19. Reason provided if some questions not addressed	0%

Item from the Evaluation Quality Checklist	Rated positively
<b>Findings</b>	
20. Social science methods were used	95%
21. Transparent connection to data source(s)	77%
22. Findings supported by data from range of methods	88%
23. Findings distinct from conclusions/recommendations	72%
24. Findings are precise	86%
25. Unplanned/unanticipated results were addressed	15%
26. Alternative possible causes were addressed	13%
27. Person-focused findings sex-disaggregated at all levels	55%
28. Report discussed gender differential access/benefits	59%
<b>Recommendations</b>	
29. Recs limited to recommendations, free of findings	86%
30. Recs specific about what is to be done	89%
31. Recs specific about who should take action	66%
32. Recs clearly supported by findings	72%
<b>Annexes</b>	
33. SOW included as an annex	80%
34. List of sources included as annex	71%
35. Data collection instruments included as annex	88%
36. All data collection instruments included	79%
37. Statements of Differences included as annex	3%
<b>Evaluation Data Warehousing</b>	
38. Report explains how data will transfer to USAID	2%
<b>Link to Evaluation Policy Quality Standards</b>	
39. Evaluation SOW includes Evaluation Policy Appendix I	44%
<b>Additional Questions About Basic Evaluation Characteristics</b>	
40. Evaluation includes table of contents	100%
41. Report includes glossary/list of acronyms	100%
42. Report well-written	89%
43. Report is well-organized	86%
44. Published date on report	99%
45. Team leader named in report	66%

## Executive Summary

Of the 177 evaluations in this assessment, all but one included an executive summary. Of these, 86 percent included critical report elements such as the evaluation purpose, questions, background information, methods, study limitations, findings, and recommendations.

<b>Executive Summary</b>	
1. Executive summary mirrors critical report elements	86%

## Introductory Information

Evaluations did a good job<sup>13</sup> of describing the activity, with 92 percent of evaluations scoring positively. However, they only did a fair job (77 percent) of describing the activity's theory of change or development hypothesis.

Activity/Program/Project Background	
2. Project characteristics described	92%
3. Project "theory of change" described	77%

## Purpose of the Evaluation

Evaluations also did a good job describing the evaluation's management purpose, with 93 percent of evaluations detailing the specific decisions or actions the evaluation is expected to inform.

Evaluation Purpose	
4. Management purpose described	93%

## Evaluation Questions

Almost all evaluations had questions that were linked to the evaluation's purpose. However, the link between the questions in the evaluation SOW and the report was less clear. The questions in the body of the report exactly matched the SOW in 68 percent of evaluations. Differences between the report and SOW are not necessarily a sign of poor quality, as it may be acceptable for an evaluation team to revise the questions in consultation with the evaluation commissioner during the inception phase. However, in cases where the questions did not match, only 14 percent clearly stated that they had received approval to change the questions.

Evaluation Questions	
5. Questions were linked to purpose	99%
6. Questions in report same as in SOW	68%
7. Written approval for question changes documented	14%

Evaluations did a good job of being responsive to evaluation questions. Eighty-eight percent of evaluations were structured to respond to evaluation questions as opposed to themes or some other structure. Evaluations adequately addressed their evaluation questions in the body of the report, with 91 percent responding to all questions. However, in the nine percent of cases that did not answer all questions, none explained why the questions were not addressed.

---

<sup>13</sup> As outlined in the introduction to this section, "good" is defined as 80 percent or higher of evaluations rating positively on the factor; "fair" is 50 to 79 percent, "marginal" is 25 to 49 percent, and "weak" is less than 25 percent.

<b>Responsiveness to Evaluation Questions</b>	
17. Report structured to respond to questions	88%
18. Evaluation questions addressed in body of report	91%
19. Reason provided if some questions not addressed	0%

Additional analysis on evaluation questions is provided in Section 3.4.2. The assessment team also extracted the text of the evaluation questions for future analysis. They are included in the spreadsheet deliverable.

### Team Composition

All factors related to team composition ranked as “fair”. For an evaluation to be rated positively for an external team leader (65 percent), evaluation specialist (55 percent), and local team members (58 percent), the evaluation had to explicitly state these facts. In discussions with the assessment team, many evaluations that were rated negatively due to lack of clarity. It is the assessment team’s opinion that for factors 12 through 14, the actual number of external team leaders, evaluation specialists, and local team members may be higher than what was clearly documented in the evaluation reports.

Only half (50 percent) of evaluations documented that team members had signed conflict of interest forms, either by stating so or including them as an annex. All team members are required to sign a conflict of interest statement according to the USAID Evaluation Policy.

<b>Team Composition</b>	
12. External team leader	65%
13. Evaluation team included an evaluation specialist	55%
14. Evaluation team included local members	58%
15. Conflict of interest forms signed	50%

### Evaluation Methodology

Evaluations did a good job of describing their data collection methods (100 percent) and to a lesser extent their data analysis methods (80 percent). However, they only did a fair job of linking these methods to their evaluation questions (62 and 60 percent respectively). The types of data collection and analysis methods, their demonstrated use, and method of data collection (in-person or virtual) are detailed above in sections 3.4.5. and 3.4.6.

While not covered on the original Evaluation Quality Checklist, this assessment also looked at whether the evaluation discussed any gender-sensitive data approaches to data collection and found that this area ranked as “weak”, with only one third of evaluations documenting this practice.

<b>Methodology</b>	
8. Data collection methods described	100%
9. Data collection methods linked to questions	62%
-- . Gender sensitive data collection methods	33%
10. Data analysis methods described	80%
11. Data analysis methods linked to questions	60%

### Study Limitations

Evaluations did a good job of including the limitations of their evaluations, with 95 percent of evaluations including them in the body of the report before the findings section.

<b>Study Limitations</b>	
16. Study limitations included	95%

### Evaluation Findings

In the presentation of findings, evaluations did a good job at demonstrating use of social science methods (95 percent), utilizing findings from multiple data collection methods (88 percent), and being precise in findings, rather than using language such as some, many, most (86 percent). However, evaluations only did a fair job in transparently connecting findings and their data sources (77 percent) and keeping their findings distinct from any conclusions and recommendations (72 percent).

Very few evaluations addressed any unplanned or unanticipated results (15 percent) or alternative possible causes of outcomes (13 percent). While these factors are rated as “weak”, these are areas that not all evaluations might be expected to address.

In terms of gender considerations, evaluations included sex-disaggregated data at all levels where findings were person-focused (55 percent) and addressed whether access/participation and/or outcomes/benefits were different for men and women (59 percent).

<b>Findings</b>	
20. Social science methods were used	95%
21. Transparent connection to data source(s)	77%
22. Findings supported by data from range of methods	88%
23. Findings distinct from conclusions/recommendations	72%
24. Findings are precise	86%
25. Unplanned/unanticipated results were addressed	15%
26. Alternative possible causes were addressed	13%
27. Person-focused findings sex-disaggregated at all levels	55%
28. Report discussed gender differential access/benefits	59%

## Evaluation Recommendations

Evaluations did a good job of limiting their presentation of recommendations to just recommendations, defined as free from repetition of information already presented or new findings not previously revealed (86 percent). In terms of specificity, evaluations did a good job of being specific about what is to be done (89 percent) but were less clear about who the recommendations were directed towards (66 percent). Evaluations did a fair job of creating a clear path between their findings, conclusions, and recommendations (72 percent).

Recommendations	
29. Recs limited to recommendations, free of findings	86%
30. Recs specific about what is to be done	89%
31. Recs specific about who should take action	66%
32. Recs clearly supported by findings	72%

The assessment team extracted all the evaluation recommendations for future analysis. They are included in the spreadsheet deliverable.

## Annexes

Evaluations did a good job of including the evaluation SOW as an annex (80 percent). They did a fair job of clearly including all sources such as list of documents reviewed and individuals or groups interviewed as an annex (71 percent). Evaluations did a good job of including their data collection instruments (88 percent), but only a fair job at including an instrument for every data collection method (79 percent).

Only three percent of evaluations contained one or more statements of difference. While this puts this factor in the “weak” category, this is not something that would be expected to be found in all evaluations. A statement of difference would only be included if one was written by a team member, implementing partner, Mission, etc. in response to the evaluation.

Annexes	
33. SOW included as an annex	80%
34. List of sources included as annex	71%
35. Data collection instruments included as annex	88%
36. All data collection instruments included	79%
37. Statements of Differences included as annex	3%

## Data Reporting

Only two percent of evaluations explained how data would be transferred to USAID. Datasets are now supposed to be submitted to the Development Data Library through the Program Office, which did not exist when the Evaluation Quality Checklist was developed. It is the assessment team’s opinion that this factor is no longer relevant.

Evaluation Data Warehousing	
38. Report explains how data will transfer to USAID	2%

## Evaluation Quality Standards

Only 44 percent of evaluation SOWs include the mandatory language on evaluation report quality, resulting in a ranking of “marginal” on this factor. This text was Appendix I of the 2011 USAID evaluation policy and was later modified and integrated into the ADS as a mandatory reference.<sup>14</sup> Either version of the text as well as reference to Evaluation Policy Appendix I or the ADS mandatory reference were accepted as a positive rating in this assessment.

Link to Evaluation Policy Quality Standards	
39. Evaluation SOW includes Evaluation Policy Appendix I	44%

## Additional Evaluation Characteristics

In general, evaluations did a good job on the quality factors related to other basic characteristics. All evaluations included a table of contents and a glossary or list of acronyms and almost all included the date that the report was published (99 percent). The assessment team found that the reports were mostly well written (89 percent) and well organized (86 percent). However, only 66 percent of evaluations clearly named the team leader.

Additional Questions About Basic Evaluation Characteristics	
40. Evaluation includes table of contents	100%
41. Report includes glossary/list of acronyms	100%
42. Report well-written	89%
43. Report is well-organized	86%
44. Published date on report	99%
45. Team leader named in report	66%

<sup>14</sup> <https://www.usaid.gov/sites/default/files/documents/201maa.pdf>

## 4.2. Conclusions on the Quality of Evaluations

The overall quality of Africa evaluations has improved according to the Evaluation Quality Checklist. The evaluation quality scores increased as compared to previous studies and are clustered on the top end of the 10-point quality scale.

Table 6 ranks all the quality factors from most positive to least. The assessment team used these rankings to conclude where the greatest opportunities for improvement lie.

Areas of improvement can be grouped into structural/clarity issues and methodology issues. While the responsibility for most of these factors sits primarily with the evaluation team, they are also affected by how USAID writes SOWs, manages the evaluation process, and reviews draft evaluation reports. The assessment team concludes that these areas are where USAID could focus efforts to improve evaluation quality.

### Structural and clarity areas for improvement

Easier areas for improvement for evaluation teams may be those factors related to the structure or basic clarity of the evaluation report. These are listed below from the least to most positive result in this assessment and refer to the associated quality factor number. The assessment team concludes that evaluation quality could be improved by targeting the following areas.

- Ensuring that the questions in the evaluation report and the SOW match (6), and if they do not clearly state that USAID approved the change (7)
- Signing and include the mandatory conflict of interest forms for evaluation team members in the report (15)
- Describing the evaluation team clearly – whether the team leader is external (12), and if the team included an evaluation specialist (13) or local team members (14)
- Defining which data collection (9) and data analysis methods (11) are used to answer each evaluation question
- Writing recommendations so that it is clear who they are targeted towards (31)

Table 6: Evaluation quality factor rankings

Factor from the Evaluation Quality Checklist	Rated positively
8. Data collection methods described	100%
5. Questions were linked to purpose	99%
20. Social science methods were used	95%
16. Study limitations included	95%
4. Management purpose described	93%
2. Project characteristics described	92%
18. Evaluation questions addressed in body of report	91%
30. Recs specific about what is to be done	89%
17. Report structured to respond to questions	88%
35. Data collection instruments included as annex	88%
22. Findings supported by data from range of methods	88%
29. Recs limited to recommendations, free of findings	86%
24. Findings are precise	86%
1. Executive summary mirrors critical report elements	86%
10. Data analysis methods described	80%
33. SOW included as an annex	80%
36. All data collection instruments included	79%
21. Transparent connection to data source(s)	77%
3. Project "theory of change" described	77%
32. Recs clearly supported by findings	72%
23. Findings distinct from conclusions/recommendations	72%
34. List of sources included as annex	71%
6. Questions in report same as in SOW	68%
31. Recs specific about who should take action	66%
12. External team leader	65%
9. Data collection methods linked to questions	62%
11. Data analysis methods linked to questions	60%
28. Report discussed gender differential access/benefits	59%
14. Evaluation team included local members	58%
13. Evaluation team included an evaluation specialist	55%
27. Person-focused findings sex-disaggregated at all levels	55%
15. Conflict of interest forms signed	50%
39. Evaluation SOW includes Evaluation Policy Appendix I	44%
25. Unplanned/unanticipated results were addressed	15%
7. Written approval for question changes documented	14%
26. Alternative possible causes were addressed	13%
37. Statements of Differences included as annex	3%
38. Report explains how data will transfer to USAID	2%
19. Reason provided if some questions not addressed	0%

- Including sources (34) and all data collection instruments (36) as annexes
- Describing the theory of change, results framework, etc. in the beginning of the report (3)

Additional areas for structural improvement by USAID include:

- Incorporating the required language from the ADS mandatory reference *Criteria to Ensure the Quality of the Evaluation Report* in evaluation of SOWs (39)

## Methodological areas for improvement

Other areas for improvement may be more complicated to achieve, such as those where evaluation methodology or the implementation of the methodology needs to be addressed by the evaluation team. The points below include additional conclusions based on the assessment team's interpretation of the underlying causes of these factors.

- Increasing the inclusion of sex-disaggregated data (27) and discussion of gender differential access and results (28). This may require collecting data differently to ensure that findings can be disaggregated by sex at all levels and that data are collected on gender differentials. If the data are being collected, evaluation teams may need to improve their analysis and presentation of this information in the evaluation reports.
- Including an evaluation specialist (13) and local team members (14). These factors are included both here and under the structural/clarity category above, as it is unclear if evaluation teams are not including these types of team members, or simply not specifying their inclusion in the report.
- Distinguishing clearly between findings, conclusions, and recommendations throughout the report (23). This includes not only following a structure that clearly delineates between findings, conclusions, and recommendations, but also ensuring that no conclusions or recommendations are written into the findings section.
- Making a clear connection between the findings, conclusions, and recommendations so that the reader understands the basis of all recommendations (32). Successful examples of this quality factor included providing a findings, conclusions, recommendations table (also called a Getting to Answers matrix) as an annex; including recommendations after each evaluation question's findings and conclusions section as opposed to at the end of the report; and clearly linking the recommendations to the questions in a standalone recommendations section.
- Writing transparently so that readers can follow a clear connection between findings and their sources (21). In cases where these connections are unclear, it can also be difficult for a reader to understand if something is a finding (a statement of fact from data collected) or a conclusion (the evaluation team's interpretation of the findings), which also affects quality factor 23.

## Meta-evaluation tools

The assessment team's final conclusion is on the Evaluation Quality Checklist tool itself. The tool was designed in 2012 structured around USAID's 2011 Evaluation Policy. Evaluations have shown great improvements in the past decade to achieving the principles outlined in the policy and its successors. This trend has been documented by this assessment, as well as those in the E3 evaluation synthesis studies. The tool remains a valid one to track improvement over time and to highlight opportunities for improvement, such as those listed above. But with scores converging at the top end of the 10-point scale, it may be time to broaden the definition of quality and for future meta-evaluations to delve into a more detailed analysis to advance learning and push the quality of evaluations to a higher level.

## 5. CONTENT OF EVALUATIONS

*Question 2: What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?*

To answer this question, the assessment team developed the descriptive data and content tool to extract programmatic lessons, described above in Section 2.2.1. Five key thematic topics were analyzed by the Senior Evaluation Specialists. Their findings and conclusions are presented in Section 5.2 below. The other topics were not analyzed for this report due to time constraints of the contract.

### 5.1. Coverage of Thematic Topics

As shown in the table below, the coverage of thematic topics ranged from 85 percent (sustainability) to 12 percent (youth topics). COVID-19 was addressed in 12 percent of all evaluations. However, one would not expect to see references to COVID-19 in the 2017-2019 evaluations before the pandemic began. COVID-19 was addressed in 51 percent of evaluations published in 2020 and 2021. Findings on the top five thematic topics are discussed in the following section.

*Table 7: Coverage of thematic topics in evaluations*

Thematic topic	Percent of evaluations that addressed the topic
Sustainability	85%
Monitoring systems, learning, and adaptive management	72%
Government engagement	66%
Beneficiary selection and targeting	56%
Theory of Change	48%
Performance target setting	47%
Innovation	23%
Unanticipated results	15%
Youth capacity building	12%
Youth civic engagement and leadership	12%
COVID-19	12%

## 5.2. Findings on Key Thematic Topics

### 5.2.1. Sustainability

#### Overview

Sustainability is essentially concerned with the outcomes of a project and the interventions that support the sustainability of those outcomes. This includes the extent to which those outcomes have positively influenced and empowered the project's intended beneficiaries, their communities, and systems that affect them, to meet development needs (usually, but not always) defined in a Theory of Change within a project's Statement of Work (SOW). With this in mind, clarification of the following issues and questions assumed a significant role in highlighting what and how data extracted from the evaluations should be interpreted.

The review of sustainability assessed whether a sustainability plan was in place at either the start or the end of a project. In principle, such plans would encompass the measures taken that contribute to the flow of benefits to beneficiaries, and an exit strategy designed to ensure continuity once the external funding and support end. In addition, the review considered whether sustainability had been defined at all (in the presence or absence of a sustainability plan) as the concept can have a different meaning for project management and intended beneficiaries. Finally, the review considered whether sustainability was reflected differently in mid-term and final evaluations and if so, what lessons were learned for project design, implementation, and performance.

In reviewing all 177 evaluations, it is clear that sustainability is dealt with differently by projects and evaluation teams depending on the emphasis given to it by USAID, project management and/or the evaluation team in pursuit of their respective objectives. The team's analysis revealed the following related to sustainability, as shown in Figure 22:

- i) One hundred and five of the 177 evaluations (59 percent) addressed the issue of sustainable interventions. Thirty-two percent of those were mid-term evaluations and 68 percent final evaluations.
- ii) One hundred of the 177 evaluations (56 percent) addressed the issue of sustainable outcomes. Forty-two percent of those were mid-term evaluations and 58 percent final evaluations.

These statistics indicate that more emphasis was placed on sustainability in final evaluations for both outputs and outcomes than in mid-term evaluations. It is useful to highlight examples of how evaluations dealt with sustainability. Table 8 below provides examples on how the sustainability of outcomes and interventions was discussed in mid-term and final evaluations.

Figure 22: Sustainability in evaluation reports

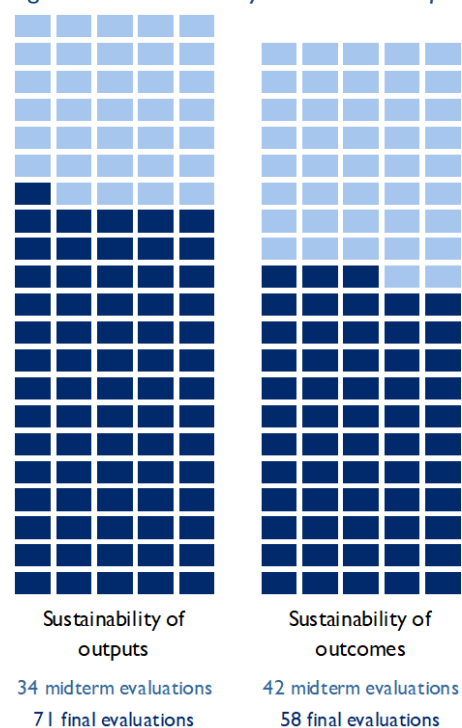


Table 8: Examples of evaluations addressing sustainability

	Sustainability of Outcomes	Sustainability of Interventions
<b>Mid-term evaluation</b>	<p><i>Feed the Future Nigeria Livelihoods Project</i></p> <p>“The Activity has helped beneficiaries to adopt important coping skills, behaviors, and practices that should support sustainability. For example, belonging to collective groups such as PGs and SILCs can mean less risk, better support during emergencies, access to better information, and personal support from group members; all these help HHs reduce the myriad risks borne by changes in the economic climate”. [Positive assessment]</p>	<p><i>Enhancing the Status of Human Rights Protection and Systems of Good Governance in Ethiopia</i></p> <p>“A majority of justice sector trainees interviewed doubted the sustainability of the training, study tour, and workshop achievements. The HRPS Activity did carry out awareness campaigns, including trainings, market forums, community conversations and testimonial events. However, Justice for All-Prison Fellowship Ethiopia interventions in the target districts have tended to be one-off, which compromises sustainability”. [Negative assessment]</p>
<b>Final evaluation</b>	<p><i>Evaluation of the Pamoja Twajenga - Capacity Development for Partners of Accountability Program (Tanzania)</i></p> <p>“USAID staff noted that awards to the ES and two direct grantees (TACOSODE and NACOPHA) continued after PT closed. The three staff that continued to manage these awards noted that these partners continued to produce high quality reporting documents as they had with PT, suggesting that the achievements in stronger reporting from PT carried over and had been sustained”. [Positive assessment]</p>	<p><i>End-term performance Evaluation for Youth Potential (Ethiopia)</i></p> <p>“Although the BPY small grants initiative demonstrated the possibility of TVET capacity building, its limited scale and primary focus on providing scholarships will likely limit the sustainability of this intervention component”. [Negative assessment]</p>

*Evaluation Sustainability “Success Story”: USAID/Tanzania SAUTI Project*

One success story articulates how the project positively contributed to the continuity of intervention activities, as noted in the evaluation report findings, and substantiated by a FGD respondent regarding the outcome of those activities as it related to her and her community. In this instance, with reference to the Tanzania Mid-term Evaluation of the SAUTI Project, the project had a sustainability plan at the start of activities and delivered initiatives designed to contribute to post-project sustainability. This is significant in that it draws attention to the project’s handling of sustainability planning as well as the evaluation team’s research and analysis of it as elaborated in the evaluation report. Given the sensitive and important nature of the sector under review (HIV prevention) as it pertains to the well-being of vulnerable populations, it is encouraging to note that both project management (and by default USAID’s role in the project design) and intended target audiences are aligned with a

**USAID/Tanzania SAUTI Project Mid-term Evaluation**  
*April 2019*

Quote from an FDG participant: “Yes, the changes will prevail because the changes will remain there. We are people of the same generation/age throughout these five years of the Project’s education. The younger ones who receive this education will be like shield of the SAUTI Project. Others when they give birth to their offsprings they will educate them with education they receive from SAUTI Project.”

common purpose in mind; that is, to contribute to and participate in behavioral change that will lead to the desired results defined in the project's SOW regarding sustainability.

Furthermore, this evaluation report raised some outstanding issues in its recommendations that needed to be addressed and proposed. Practical suggestions for the remainder of the project aimed at enhancing the prospects for longer term continuity.

“Sauti already manages sub-recipient performance very closely when it comes to achieving targets. There is some evidence from the field that intervention quality, however, is not consistent across all sub-recipients. Revisiting and investing more in the QI strategy, with a view to strengthening Project fidelity, will serve to strengthen outcomes, beyond target achievement.”

### **Lessons for Project Design, Implementation, and Performance**

The design of a project as elaborated by USAID, usually in a SOW, where the purpose is often defined based on a Theory of Change, development hypothesis or assumption, highlights the modus operandi for activity implementation and delivery of anticipated results. This includes the extent to which achievements gained can continue after the project ends and where intended beneficiaries and stakeholders have the means to embracing and taking ownership of positive changes in their favor, leading to long term improvement in their livelihoods. Analysis of the evaluations revealed that beneficiary and stakeholder perceptions of sustainability can differ depending on whether they were mid-term or final evaluations.

Only three midterm evaluations mentioned projects having a sustainability plan at the beginning of the activity. For final evaluations, 10 projects had a sustainability plan at the start and 6 at the end of the project, with 3 having both. Although this is a rather small number relative to the total 177 evaluations reviewed, the evidence suggests that inclusion of a sustainability plan is not necessarily a pre-requisite for sustainability of outcomes post-project, as discussed below regarding mid-term and final evaluation reports. In addition, the absence of a sustainability plan at the beginning or end of a project does not automatically exclude the possibility that desired change can occur, particularly if intended beneficiaries and/or stakeholders perceive project interventions and outcomes as having had a positive influence that could lead to the likelihood of continuity and improvement in their daily lives post-project.

In this respect, for mid-term evaluations, two things can be observed. First, projects that had a sustainability plan in place were viewed positively by evaluation teams based on their findings, where emphasis appears to have been on reporting on project management's close collaboration with beneficiaries and stakeholders who acknowledged the rationale of having a sustainability plan. Second, where no plans existed, there was a tendency for the majority of evaluations to either make no mention of sustainability in their findings or in a few cases simply refer to it in a recommendation though not always directly connected with a particular conclusion.

Regarding final evaluations, where a sustainability plan was referred to in their reports, there was a tendency for findings and conclusions to focus on the fact that at project end, and according to various respondents, external factors could affect continuity of any project achievements – political, economic, environmental, among others. Where no sustainability plan existed, evaluation reports either did not refer to it (as was the case for mid-term evaluation reports) or focused on two specific issues. First, reasons why sustainability was unlikely post-project, often solely based on feedback from interviewee respondents and second, reasons why continuity may occur if measures already in place were not adversely affected by unforeseen circumstances or matters beyond the control or influence of the intended beneficiaries or stakeholders. For final evaluations it was not possible to comment on

attribution issues related to sustainability as this was outside the scope of the study and which could also have been influenced by external factors.

## **Conclusions**

In summary, the issue of sustainability is complex and often clearly misunderstood regarding how it should be considered in terms of how projects or beneficiaries relate to it. In most cases, as there is no way of knowing whether continuity of outcomes prevails post-project unless there is follow-up by USAID, most evaluations essentially offered recommendations that they believed should be undertaken by the project or other parties to ensure sustainability. Furthermore, in both mid-term and final evaluations, there was little distinction drawn between interventions and outcomes though they are inextricably linked. In terms of project design, though more a recommendation than a lesson learned, projects would be better served if sustainability plans were embedded in Theories of Change or equivalent.

The fact that sustainability was addressed one way or another in most of the evaluations, suggests that evaluations did consider it, though the extent to which depended on whether a question on sustainability was included in the evaluation's SOW or sustainability plans formed part of the project design. If evaluations are to look at sustainability at least one evaluation question should directly ask what the evaluation found actions taken that contribute towards the likelihood of outcomes continuing after the project ends.

Defining sustainability during project design stage supports the development of meaningful and realistic findings, conclusions and recommendations at the evaluation stage. This would also help fulfill the management purpose of the evaluation if a follow-on project is being considered and can be adapted to take account of sustainability recommendations.

## 5.2.2. Monitoring Systems, Learning, and Adaptive Management

### Overview

This analysis focuses on evaluation content addressing Monitoring system, Learning, and Adaptive Management. This topic was addressed in 127 of the 177 evaluations (71 percent). The aim is to provide a synthesis of trends in quality and information for improvements in design, implementation, and performance in these areas. The themes that emerged as lessons learned from the analysis point to both limitations, and good practices.

### Lessons for Project Design

The most frequently cited limitations in monitoring systems pertained to the design of indicators (49 evaluations, 39 percent). These limitations were deemed to influence the measurement of the effectiveness, efficiency, and impact of the activities/projects, as well as the ability to attribute results to interventions. Indicator design limitations included:

- i) Poor metrics/indicators (including the absence of sector-specific indicators or those that could address specific challenges and emerging opportunities, such as context indicators) and targets;
- ii) Weak indicator definitions and guidance on measurement of outcomes as described in the performance indicator reference sheets;
- iii) Failure to standardize and align indicators for reporting purposes across implementing partners;
- iv) Overreliance on output versus outcome indicators for measuring results;
- v) Too many or too few indicators;
- vi) Limited or absent indicator disaggregation (at different levels, e.g., public/private, gender, age, clinics vs clinicians, etc.);
- vii) Absent baseline values for performance indicators; and
- viii) An absence of specific indicators for measuring inequities with respect to gender, youth, and social inclusion.

Other shortcomings in the monitoring system, learning, and adaptive management observed in 10 evaluations (8 percent) were:

- i) Absent or decentralized/badly structured information management systems, which can limit ownership of data platforms and a unified data collection system;
- ii) A lack of available/accessible data (for example, at community and private sector levels for some activities/projects); and
- iii) A failure in coordinating data sharing and monitoring activities at different levels and across implementing partners.

Overall, a decentralized information management system was considered a limitation to the design of an effective monitoring system for reporting, data sharing, and data quality. For example, one evaluation attributed the success of the activity's performance monitoring to the presence of a centralized database system, stating:

“The system is linked to an analytical routine at the front end, which greatly increased access to data as staff can generate bi-variate frequency tables based on their needs and interests. Partner staff receive a copy of their dataset through the SHAREit application for further analysis.”<sup>15</sup>

However, another evaluation suggested that a centralized data management system would not necessarily be the solution in guaranteeing that data were shared across partners and stakeholders, noting:

“... despite the creation of regional databases, challenges remain in terms of country ownership and completeness of data collected. On top, not all the countries seem to be convinced by the performance of the platform and continue using a different system, which they think is better. Furthermore, some countries are not interested in sending their data on a common platform, so more sensitization is needed.”<sup>16</sup>

Additional lessons drawn from the perspective of project design in ensuring a robust monitoring, learning, and adaptive management system (observed in 27 evaluations, or 21 percent) included the importance of:

- i) Robust reporting systems and dissemination plans of quality monitoring data for adaptive management purposes;
- ii) An effective learning agenda and a plan to address knowledge gaps and use of data for adaptive management purposes; and

Securing adequate funding as a prerequisite of a sustainable M&E system. One evaluation described a best practice regarding the use of robust indicators and an effective reporting and learning plan as follows:

“The program developed a series of indicators largely focused on outcomes, as a part of their commitment to MERL, while recognizing the limitations of these measures. [The activity] also brought in the dimension of Reporting and Learning, which allowed for program iteration and adjustment based on early evidence of activity success and failure. The MERL approach, through qualitative reporting, sought to identify dimensions of behavior change and learning that better captured the enabling conditions and attributes of resilience across sectors.”<sup>17</sup>

And another evaluation depicted an inadequate monitoring, evaluation, and learning (MEL) plan with limited reporting and learning quality for adaptive management as follows:

“MEL did not appear to play a central role in technical design and program implementation. Those involved in implementation noted that, particularly under previous CoPs, the MEL team was siloed from technical teams and not sufficiently empowered....Technical teams did not

---

<sup>15</sup> Evaluation 159, p. 10

<sup>16</sup> Evaluation 46, p. 55

<sup>17</sup> Evaluation 47, p. 24

incorporate MEL and data gathering considerations at the point of task design. The secondary role played by MEL is reflected in reporting, where technical teams often report different data points to (sic) those reported in PMEP indicator data. This results in a range of data inconsistencies in quarterly and annual reports and work plans that make it difficult for USAID to have a clear picture of SATIH’s achievements. Disaggregated numbers under the components (e.g., number of linkages created) at times do not sum to their totals...This inconsistency makes it extremely difficult to track progress on specific items over time—weakening the potential for the reports to serve both as a management and communications tool.”<sup>18</sup>

Overall, relatively few evaluations (8 of 127 (6 percent)) examined closely adaptive management. Instead, adaptive management was generally discussed as part of reporting, learning, data utilization, and coordination/responsibility sharing issues and recommendations.

Some of the lessons that emerged from the evaluations for promoting adaptive management at the design stage included:

- i) Availability of a strategic approach to collaborating, learning, and adapting (with respect to use of data for decision-making), as well as stakeholder engagement and receptivity;
- ii) Presence of a research strategy to address knowledge gaps - especially with regard to scientific studies - across the different implementation settings that could support relevant management decisions;
- iii) Identification of key learning points at the outset;
- iv) Flexibility in the design of the activity/project (especially with respect to unanticipated changes in context or implementation), and the alignment of outputs and outcomes with grant/funding duration for decision-making and adaptive management.

### **Lessons for Implementation and Performance**

The analysis reflects several important lessons in the implementation and performance where factors supported an effective monitoring system, learning, and adaptive management (observed in 21 evaluations (17 percent)). These consist of:

- i) Adequate financial and human resources and capacities (also to avoid high turnover rates of trained staff due to funding allocations) in conducting monitoring activities, as well as adequate division of labor in monitoring activities across partners;
- ii) Sufficient data quality control to avoid errors in reporting and inconsistencies in data cleaning;
- iii) Consistent and robust analysis and operationalization of data at different levels including adequate analysis of trends and patterns of the monitoring data for learning and decision-making purposes, as well as tracking direct and indirect costs for cost-effectiveness analysis;

---

<sup>18</sup> Evaluation 43, p. 68

- iv) Effective supervision of monitoring activities and/or follow-up and post-activity monitoring to assess the quality of the activity or retention of key outcomes including agreement on data collection tools across partners;
- v) Tracking indicators for activities based on the availability of earmarked funds rather than on program goals and/or analysis;
- vi) Ensuring that the M&E plans are updated so that they are aligned with changes in the scale and scope of activities during the activity/project (where applicable), and avoiding frequent indicator revisions that could jeopardize the consistency/compatibility of data as well as the comparison of results over time;
- vii) Avoiding falsified or exaggerated data to meet targets.

With regard to falsified or exaggerated data, one evaluation cautioned that an “over-emphasis” on results monitoring may have unintentional negative effects depending on implementing partners. It stated:

“Klls held in June 2018, after the REEEP had closed, found that at least one private-sector partner said that his company felt “policed” by the REEEP staff and pressured to provide numbers for their monitoring. Thus, numbers given may be skewed higher than the reality on the ground warranted. This does not invalidate the results shown in the indicators but does provide a cautionary note to take steps to validate data while not overburdening the suppliers of that data.”<sup>19</sup>

As part of lessons learned in the implementation and performance in monitoring, learning, and adaptive management, another evaluation concluded that:

“The expectation for innovator companies to apply the same systems and processes for monitoring and evaluation (M&E), documentation and reporting as donor projects is not only unrealistic but also unwise.... The notion of innovator companies reporting on milestone achievements that in turn trigger funding reflects an inherent conflict of interest.”<sup>20</sup>

The evaluations discerned some good examples of implementation and performance as demonstrated by the following statements:

“After a slow start in Year One, the activity significantly strengthened its Monitoring and Evaluation department.... Monitoring and evaluation staff are providing much ‘value added’ in terms of thinking about how to measure performance of various outputs, how to supervise on-going activities better, how to ensure that project activities are on track.”<sup>21</sup>

“The team has observed that [the Monitoring, Evaluation, Reporting and Learning, MERL] system has inbuilt checks for data completeness and consistency. The data are then encoded and analyzed by the [activity] data management person using the MERL system, and reports are sent to all stakeholders including local governments and other partners.”<sup>22</sup>

---

<sup>19</sup> Evaluation 105, p. 36

<sup>20</sup> Evaluation 94, p. 2

<sup>21</sup> Evaluation 22, p. 24

<sup>22</sup> Evaluation 148, p. 36

One evaluation addressed lessons learned for adaptive management by exploring ways the activity was changed throughout implementation to address shortcomings or changing circumstances.<sup>23</sup> Two evaluations' evaluation described good practices in adoptive management, respectively, as:

“Adaptive management is a key capacity that has increased. Adaptive management was the focus of much of AWARD’s training and learning by doing approach. This included engaging participants directly in taking a systems approach to problem or situation analysis, including planning climate change adaptation measures. Multi-sector problem analysis, multi-stakeholder participation and multi-intervention approaches were central to new and effective ways of working, building on the capacities and interests of stakeholders. Strengthening these capacities among stakeholders helps lay the foundation for sustainable program activities.”<sup>24</sup>

“The outstanding feature of Sauti implementation that is demonstrably linked to strong performance is the practice of intensive, data-driven adaptive management. Sauti amasses, analyses and uses data to: (i) routinely track performance against targets on a daily basis, and manage the performance of sub-recipients towards achieving those targets; and (ii) identify emerging challenges to performance, diagnose the causes of faltering performance, and develop, implement and track the efficacy of solutions.”<sup>25</sup>

## Conclusions

The majority (71 percent) of the evaluations reviewed discussed some aspect of the monitoring system, learning, and adaptive management components of the USAID activities/projects in the Africa region - although with varying degrees of focus on each sub-component (i.e., monitoring system or learning or adaptive management). Of these, only 6 percent specifically addressed adaptive management issues. Whilst the analysis depicts instances of good practices, evaluations revealed the need for greater diligence in optimizing the design, implementation, and performance of the monitoring system, learning, and adaptive management components of the interventions.

---

<sup>23</sup> Evaluation I

<sup>24</sup> Evaluation 44, p. 32

<sup>25</sup> Evaluation 77, p. 17

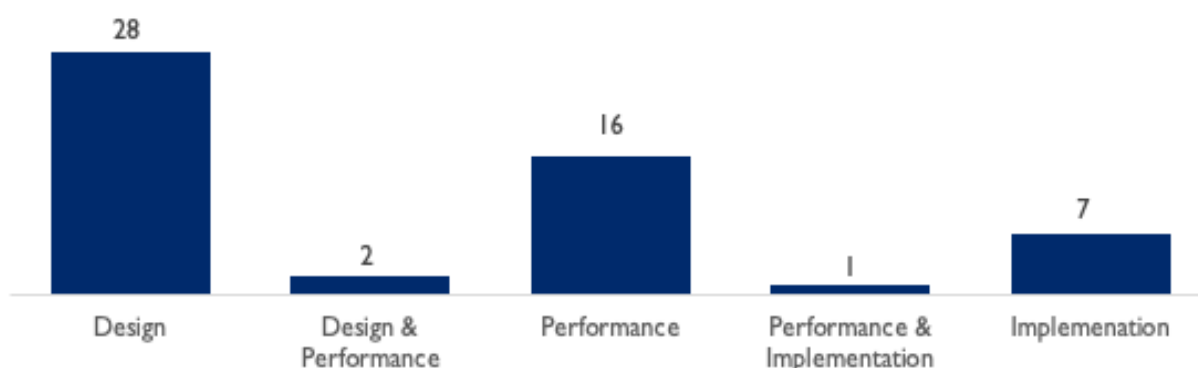
### 5.2.3. Theory of Change and Assumptions

#### Overview

This section focuses on findings relating to how the included evaluations addressed the theory of change (ToC) that informed the interventions. Within USAID, a ToC is generally articulated in a Results Framework, which defines the results necessary to achieve a strategic objective and their causal relationships and underlying assumptions. The framework establishes the organizing basis for measuring, analyzing, and reporting results of the operating unit.

As shown in Figure 23, 54 evaluations (31 percent) commented on successes or failures related to the activity's ToCs. The ToCs were variously presented/framed in the evaluation reports including presenting them as a framework, development hypothesis, or logic models. Twenty-eight evaluations discussed ToC design, 7 discussed issues around implementation, and 16 discussed issues around performance. Two evaluations commented on both the design and performance of the ToCs and one evaluation commented on both the performance and implementation fidelity of the ToC. Eight evaluations included a specific question on assessing the robustness or validity of the activity's ToC.<sup>26</sup>

Figure 23: Coverage of ToC lessons in evaluations



#### Lessons for Project Design

The key themes across the 28 evaluations focusing on ToC design were the need to use evidence to inform the design of ToCs, ensuring active stakeholder and beneficiaries' involvement in the design of ToCs, taking into account project context, making explicit assumptions around the ToC, and alignment of the ToC to the needs of project beneficiaries. Across the evaluations, a recurrent theme was the need for evidence-informed ToCs, and the involvement of stakeholders in the design process. In terms of incorporating the needs of the beneficiaries/stakeholders into the ToC, one evaluation noted as follows:

“The Amalima project should have organized community needs assessments in the planning formulation phase, as well as developed a detailed, sectoral TOC from which to justify and formulate specific project activities.”<sup>27</sup>

In another evaluation, stakeholders' participation in the design of the ToC was found to have aided the identification of assumptions and the development of an appropriate program results chain.<sup>28</sup> This clearly

<sup>26</sup> Evaluations 23, 45, 83, 116, 134, 163, 171, and 174.

<sup>27</sup> Evaluation 40

<sup>28</sup> Evaluation 7

highlights the utility of stakeholder involvement in the development of ToCs. The key lesson is that stakeholders can be engaged in developing the ToC from the beginning, or an initial ToC can serve as a tool to engage relevant stakeholders in discussion and refinement of the pathways.

Some evaluations noted that the ToCs were well thought through and provided a clear results chain for the interventions. One evaluation noted as follows:

“All components of the project fed into the Theory of Change in that the project supported the communities to develop management plans for their fisheries assets based on their shared experiences and advice from the project.”<sup>29</sup>

### **Lessons for Implementation**

Seven evaluations reported on changes and adaptations that were made or that needed to be made to the ToC during program implementation. Four of these evaluations were midterm performance evaluations, two were final performance evaluations, and one was an impact evaluation.

One evaluation noted as follows:

“The evaluation team found that the Program’s theory-of-change should be updated based on program experience as part of the program’s adaptive management approach. For example, two shortcomings of note include the observation that there are no specific outputs/outcomes with respect to: (1) support for alternative or supplemental livelihoods; and (2) reduction in demand for and use of mangrove wood.”<sup>30</sup>

One of the key messages emerging from these evaluations is that a ToC is best when kept flexible, not prescribed. That is, a ToC should be seen as a theory of change thinking, a flexible approach to think through program logic. Across the seven evaluations, the lesson for implementation is that a ToC is most effective when applied through pre-existing processes, to support critical thinking throughout the program cycle. Another important lesson highlighted by the evaluations is that working with a ToC requires performance management strategies to accommodate uncertainty and flexibility.

### **Lessons for Performance**

The core themes from the 17 evaluations that provided information on the performance of ToCs were greater achievement of project goals and objectives, providing a foundation for program design, and a basis for implementation and evaluation. There were mixed findings in terms of the validity and functionality of the ToCs examined by the evaluations. Nine evaluations provided information on program performance in relation to their theories of change and found the ToCs contributed to better program design and improved program performance. Eight evaluations reported that the ToCs could not capture all the complexities associated with the interventions, and thus, did not result in improved program performance. In some evaluations, it was reported that the ToCs did not provide a sound framework for setting out monitoring and evaluation plans. Theories of Change can provide the basis for effective program design and evaluation and, as such, the robustness of the ToC used matters.

### **Conclusions**

Three conclusions can be drawn from the findings relating to the ToCs. First, relatively few evaluations pay attention to the program ToCs, which are themselves supposed to inform the evaluation design.

---

<sup>29</sup> Evaluation 23

<sup>30</sup> Evaluation 54

Second, ToC potential is seen as supporting programs' innovation and adaptation in response to dynamic contexts. It is thus important to make ToCs flexible and contextually responsive both in the design stage and, more importantly, in implementation and performance management. Finally, as assumptions take time to become more visible, an iterative, staged ToC process that is embedded within program learning can inspire innovations and improvements in program strategies, and strengthen the potential of programs to optimize their development outcomes.

## 5.2.4. Government Engagement

### Overview

This section provides a summary of findings related to government engagement (GE) in the selected evaluations. As USAID does not have a specific policy on GE, the assessment team used USAID's Local Capacity Development (LCD) Policy,<sup>31</sup> specifically the seven LCD principles, as a proxy to classify and synthesize ways in which projects manifested aspects of GE during the project life.

Table 8 organizes GE found in the sample evaluations into two levels, strong and weak. A strong presence of LCD principles (three or more) was found in 59 evaluations (33 percent). A weak presence of LCD principles (one or two) was found 55 evaluations (31 percent). Sixty-three evaluations (36 percent) did not document any aspects of government engagement. However, the absence of GE documentation does not mean that GE was absent during the life of the project (LOP), but rather that any GE was not reported.

### USAID's Seven Principles to Guide Effective Local Capacity Development

1. Start with the local system.
2. Develop diverse capacities through diverse approaches.
3. Align capacity development with local priorities.
4. Appreciate and build on existing capacities.
5. Be mindful of and mitigate the unintended consequences of our capacity development.
6. Practice mutuality with local actors.
7. Measure performance improvement in collaboration with local actors.

*Source: USAID LCD Policy Version 8, 2021*

*Table 9: Levels of government engagement in evaluations*

GE Level	USAID LCD Principles	Number of Evaluations	Percent of Evaluations	Principles Documented						
				1	2	3	4	5	6	7
Strong	3 or more LCD Principles present	59	33%	X	X	X	X		X	X
Weak	1-2 LCD Principles present	55	31%	X	X	X				X
Unknown	N/A	63	36%							

### Lessons for Strong Government Engagement

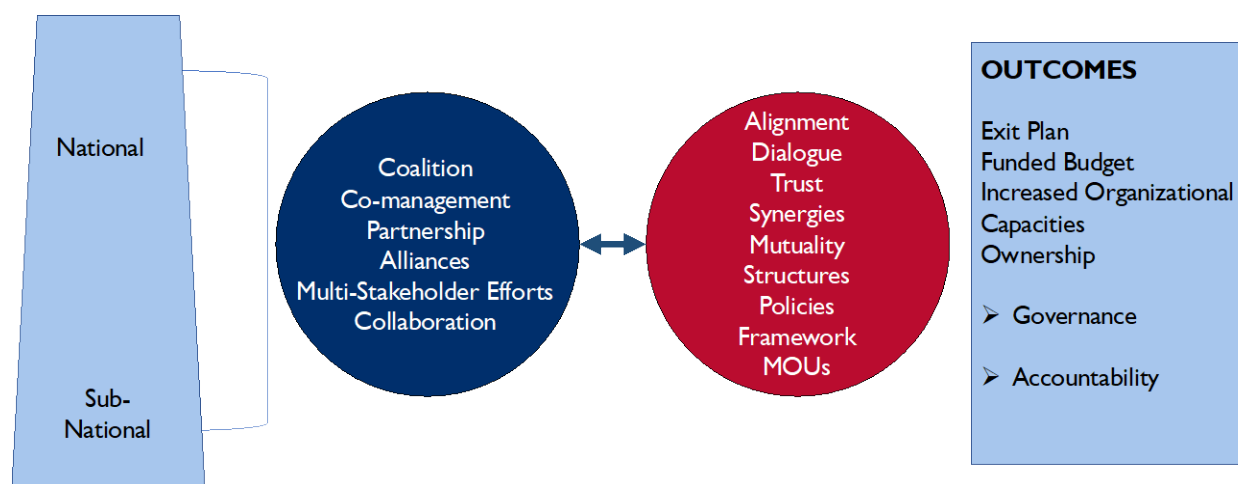
The evaluations with strong GE have at least three LCD principles present. The essential three principles were 1 or 3; 2 or 4; and 6 or 7. How the principles were operationalized depended on existing project context, government actors, and implementing partners. Evaluations with high GE seemed to have strong collaborative relationships with the IP and/or other stakeholders. In these evaluations, there were reports of: trust between government, project and/or other actors; local

<sup>31</sup> USAID LCD Policy. (2021) Version 8.

(government) appreciation for the Activity; and project ownership. These evaluations also reported regular meetings and information sharing among actors throughout the LOP, including “co-managing” field activities with government actors.

Figure 24 illustrates some key GE elements and LCD Principles documented in the evaluations, their processes (dialogue, trust), products (MOUs, frameworks), and outcomes (exit plan, budget for activities). In almost all cases, implementers interacted with the government stakeholders at all levels. Principle 5, “be mindful of and mitigate unintended consequences of our capacity development,” was not found in any evaluations.

Figure 24: Principles for effective local capacity development and strong government engagement



Strong GE was found in a myriad of implementing organizations, sectors, and contexts, for example:

- Evaluation 69: The project supported the creation of a coalition of 13 organizations that were invested in making the budget process more transparent and effective. The coalition maintains an active partnership with the Ministry of Economy and Finances, which previously had no interlocutor representing civil society. The multi-actor framework is a mechanism that contributes to strengthening the governance of public finances through the creation of a space for dialogue, consultation, and pooling of resources.
- Evaluation 131: As a DRG activity focused on increasing CSO capacity for advocacy, the situation has improved to one of more cordiality, mutual respect, and amicable discussion.
- Evaluation 169: The use of the inter-village committee and peace management committee mechanisms for conflict resolution by local authorities in certain villages testifies to the support and the way in which the activities’ approaches were absorbed, accepted and used. This collaboration was developed by the two IPs.
- Evaluation 12: The WASH project created management structures to support accountability of utility agencies. SOPs and consumer charters to define roles and responsibilities were developed in collaboration with stakeholders. According to the report, utility organizations have greater ownership for the operations and outcomes of service delivery.

- Evaluation 9: The project supported the Ethiopian Ministry of Finance and Economic Cooperation Public-Private Partnership (PPP) Policy, which outlined the need for private sector involvement in all sectors for economic growth, increased innovation, improved quality, and reduced delays in service delivery.
- Evaluation 67: The active participation and involvement of both national and local government bodies had a cascade with organizations working under those bodies. The high degree of coordination between the HDAK project and key government counterparts has led to the integration of the WRN! curriculum into the national technical education curriculum.
- Evaluation 7: The activity worked closely and successfully with Ministry's Evaluation Unit to develop their early grade reading assessment, including working together on tool development, implementation, analysis, and dashboard. Use of government expertise and materials was cost effective.
- Evaluation 24: The second major accomplishment was the productive partnership created and nurtured with government counterparts. The project collaborated with field extensionists, technical staff, Health Surveillance Assistants, and ministry staff in health and disaster management.
- Evaluation # 32: CSO advocacy helped push the Ministry of Women's Affairs, along with other ministries, to allocate funding for women's causes in the work planning and state budget allocation.

### **Lessons for Weak Government Engagement**

Evaluations with weak GE possessed only two LCD principles. In general projects with weak GE did not start with the local system, neither did they align with local priorities, nor appreciate local capacities. Indeed, these evaluations reported poor communication/interactions between the IP and government stakeholders which likely prohibited mutual learning and collaboration-strengthening. Consider Figure 22 as a reference, without an adequate working relationship where there are information exchanges and/or coordination efforts, there would not be trust and ownership on the part of the government stakeholder. There would also be missed opportunities to leverage local resources to strengthen or complement project strategies. With little input or engagement from government actors, the project activities were often delayed especially when government authorization was required.

There is no trend observed in terms of weak GE occurrences observed across all the sectors and countries. The following provides an illustration of the kinds of challenges reported:

- Evaluation 34: The project had “limited interface” with key government actors which demotivated government staff who were supposed to monitor project implementation. Monitoring efforts suffered due to the weak connection.
- Evaluation 25: Although the project engaged with both national and local government levels it was not always entirely successful.
- Evaluation 77: “The attitude of the government of Tanzania toward homosexuality continues to constrain implementation.”

## Conclusions

If government capacity building/partnership is not directly asked about in an evaluation question or is an integral part of the activity's intervention, evaluations often do not document and provide this information in the report. This does not mean that some aspects of GE were not present during the LOP. Understanding how and why government engagement happens is important as presented in the previous sections, especially as USAID is beginning to recognize that sustainable transformations require systems change.<sup>32</sup> Systems change requires project implementers to collaborate and learn from each other and adapt to the internal and external forces. Dismissing or missing opportunities to work and learn from local partners may bring about unintended (latent) consequences.<sup>33</sup>

### 5.2.5. Beneficiary Selection and Targeting

#### Overview

Of the 177 evaluations reviewed, 99 addressed beneficiary selection and targeting (approximately 56 percent), and/or included data on how beneficiaries were identified to take part in the evaluation process. Most of these 99 evaluations described specific groups and/or the need to include others, and at times there were some insights on the extent to which a targeted group appeared relevant.

There were few differences identified amongst the different kinds of evaluation, (i.e., midterm vs. final, impact vs. performance, activity type), in terms of what was discussed. Most discussions provided a description of beneficiaries and/or targets, not an assessment of how they were selected or a recommendation on how to shift or otherwise change the beneficiary group. The few that did offer insights are discussed below.

#### Lessons for Project Design

There were few recurrent themes or lessons learned. On project design, evaluations tended to indicate that USAID projects/programs addressed or included women, youth and other groups who are often marginalized in some manner. Only a few evaluations identified significant issues with targeting or selecting appropriate beneficiary groups. One example discussed how the intended group was not directly/sufficiently engaged: "SEUH's design also did not provide for direct interventions for poor urban populations working in industrial parks but assumed that these groups were receiving services in their homes within the targeted communities."<sup>34</sup>

#### Lessons for Implementation

Nearly all 177 evaluations mentioned how beneficiaries benefitted, or did not benefit, from the intervention. A few evaluations discussed how the activity identified beneficiary groups. Some evaluations mentioned that implementation of the intervention and reaching targeted groups was often sporadic, unclear, or inconsistent.

#### Lessons for Performance

Data supporting lessons for performance were very limited. One health-focused evaluation suggested the need for better selection of women who receive health interventions. Another evaluation had a strong conclusion regarding the challenge of working in highly vulnerable areas, noting that "the

---

<sup>32</sup> USAID's LCD Policy.

<sup>33</sup> For more information on learning and Systems Thinking see Peter Senge's *The Fifth Discipline: The Art and Practice of the Learning Organization*.

<sup>34</sup> Evaluation 49

qualitative study reveals the complexities of targeting within highly vulnerable communities.”<sup>35</sup> Other studies commented on the usefulness of engaging marginalized populations, such as “active mobilization and participation of marginalized groups contributed significantly to an increase in voice and space.”<sup>36</sup>

The scarcity of data on beneficiaries and target groups does not allow for significant lessons learned on performance. The lack of attention to beneficiary selection in evaluations could result from the evaluation terms of reference not specifying the theme as a priority for USAID, or from a general perception that beneficiary selection is not a key impediment to results achievement.

## **Conclusions**

The lack of data on the topic of beneficiary selection and targeting suggests that either (1) beneficiary selection is not an identified area of focus for USAID evaluations and/or (2) monitoring or other data suggest few issues or challenges with beneficiary selection or targeting. A key lesson is that evaluations should address how and why beneficiary groups were selected and how that selection relates to the problem statement and the theory of change if this is an area of interest.

---

<sup>35</sup> Evaluation 21

<sup>36</sup> Evaluation 32

# ANNEX I: STATEMENT OF WORK

## Review and Synthesis of Evaluations in the Africa Region Conducted between 2017-2021

### I. Introduction and Background

The quality of evaluations conducted in the Africa region has gradually improved since the adoption of the Evaluation Policy in 2012. However, there are still significant quality challenges which have been documented through various evaluation quality assessments conducted by the Africa Bureau, PPL, and a few Pillar Bureaus over the years. PPL and the Africa Bureau have also compiled evaluation limitations included in the evaluation reports which also point to many factors that affect the quality of evaluations.

Several years have passed since the last evaluation quality assessment was conducted by the former Economic Growth, Education, and Environment Bureau (E3) in 2015 which covered economic growth and education evaluations, the E3/Education Office in 2016 covering education evaluations, and the former Bureau for Food Security (BFS) in 2015 covering agriculture and food security evaluations. It's therefore important to update the evaluation quality reviews that were conducted 5-6 years ago to assess whether or not there has been further progress in the improvement of the quality of AFR evaluations in order to identify evaluation quality gaps that need to be addressed and to tailor Mission evaluation capacity strengthening interventions accordingly.

The proposed assessment of evaluation quality covers a sample of evaluations conducted in the Africa Region between 2017-2021.

### II. Purpose and Use of the Assessment

#### A. Purpose

AFR expects the results of this activity will generate information that will raise awareness among Africa Mission and Africa Bureau leadership about the strengths and weaknesses of the evaluations produced by the Bureau as a whole. Information on factors affecting evaluation quality will help the Bureau and Missions to devise strategies for quality improvement. In addition, this assessment will shed some light on overall project/activity performance in the Africa Region, including some of the key factors that facilitate or inhibit performance.

During the dissemination of findings from this assessment, Missions will be encouraged to analyze other factors not covered by the assessment that influence the quality of evaluations, including internal Mission operations (evaluation planning, staff capacity, evaluation independence, evaluation management, etc), IP MEL systems, and the capacity of evaluation firms and teams conducting USAID evaluations.

#### B. Intended Uses

Africa Bureau will use the report synthesizing the findings from this assessment to determine topics on which it will develop additional guidance, products, and technical support to improve the quality of evaluation report generated by Africa evaluations.

Missions on the other hand will use the findings from this assessment to improve the quality of their evaluations by addressing the key factors that hinder the generation of high quality evaluations for evidence-based decision-making.

### **III. Existing Information Sources**

The assessment will draw from 2017-2021 evaluation reports posted in the Evaluation Registry. It will also be informed by similar assessments various Bureaus conducted in the past including:

- a. PPL Meta-Evaluation of the Quality of Evaluation Reports (2012)
- b. E3 Sectoral Evaluation Synthesis (2014) and (2015)
- c. E3/Education Assessment of the Quality of Education Evaluations (2017)
- d. Feed the Future Evaluation Synthesis (2015)
- e. Africa Bureau Review of the Quality of Evaluations in the Africa Region (2014)
- f. PPL Analysis of Evaluation Limitations included in Evaluation Reports (2018)
- g. Africa Bureau Analysis of Evaluation Limitations (2019-2020)
- h. USAID Evaluation Policy (2012) updated in 2016
- i. PPL How-to-Note on Drafting Evaluation Reports
- j. PPL Evaluation Quality Review Checklist
- k. PPL Evaluation SOW review checklist

### **IV. Key Questions of the Evaluation Quality Assessment and Synthesis**

The proposed assessment seeks to answer the following questions, consistent with the questions previous assessments focused on:

1. What do Africa evaluations (evaluations funded/commissioned by USAID Operating Units in the Africa Region) completed between 2017 and 2021 tell us about overall strengths and weaknesses in their quality? Where do the greatest opportunities for improvement lie?
2. What key programmatic lessons emerge from this review related to select aspects of project design, implementation, and performance?

### **V. Methodology**

To answer the questions above, the assessment team will review and assess the quality of evaluations completed between 2017 and 2021. The assessment will be conducted using tools/ a checklist developed by PPL and used by the E3 Bureau in previous evaluation quality assessments. The tools consist of a) a checklist comprising of a set of questions (38 in total) related to USAID evaluation quality factors/criteria to assess whether or not the report meets these factors, b) a shorter checklist comprising 10 questions to help generate an overall score of the quality of each evaluation, and c) a questionnaire to capture 4-5 aspects of program design, implementation, and performance. These will be determined by the Africa Bureau in collaboration with the assessment team. Details about these tools and other methodological details are described in the Tasks and Deliverables section below.

### **VI. Tasks and Key Deliverables**

#### **A. Tasks**

The tasks outlined in this section are illustrative and will be refined in collaboration between AFR and USDA/KNUST implementation team.

### **1. Review and Adaptation of Evaluation Quality Assessment Tools/Checklists**

Various Bureaus have developed tools to assess the quality of evaluations. In 2012, Management Systems International (MSI) under a PPL/LER contract adapted the tools that had been used in the past and created a tool/checklist to conduct a Meta-Evaluation of the Quality of Evaluation Reports. The tool, which is based on USAID evaluation quality standards and practices described in the Evaluation Policy and ADS 201, was subsequently used by MSI to conduct a Sectoral Evaluation Synthesis of E3 Bureau evaluations completed between 2013-2014 (first study) and 2015-2016 (second study). In order to assess trends in the quality of Africa Bureau evaluations since the issuance of the Evaluation Policy in 2012, we will use the same tool, with minor adaptations, to assess the quality of AFR evaluations conducted between 2017-2021. The tool consists of a checklist of 38 criteria that correlate with evaluation quality and a guide that accompanies the checklist. In addition, in its 2015 assessment, the E3 Bureau developed a content analysis tool to capture data on key project or activity themes, lessons learned, and evaluation quality factors not included in the checklist. We will adapt this tool to conduct content analysis that will feed into the final quality assessment and synthesis report.

### **2. Defining the Data Set**

The project team will work to define the universe of evaluations to be included in this report. The universe will be based on those evaluations completed in the defined time period (January 1, 2017 to December 31, 2021). AFR will compile the list of evaluations reported in the Evaluation Registry and make it available to the consultants. However, since not all Bureau/Mission evaluations are reported in the registry, the assessment team will compile another list of evaluations posted on the DEC, compare the two lists, and add any evaluations reported on the DEC that are missing in the initial list provided by the Africa Bureau. In addition, since evaluation titles can sometimes be misleading, the assessment team will ensure the evaluations included in the final list meet the USAID definition of evaluations. The assessment checklist includes key characteristics of the evaluation such as sector of the project or activity evaluated, evaluation purpose, type (performance vs impact, internal vs external), budget, timing/dates, etc. The list provided by the Africa Bureau will specify some of these criteria and the assessment team will fill in the criteria that are missing from the Africa Bureau list.

The full list of AFR evaluations completed during the above period with evaluation reports posted on the DEC or the Internet is approximately 210. Given limited time and resources, the assessment will select a representative sample of these evaluations which will be the focus of the evaluation quality review. For illustrative purposes and considering funds available, a minimum of 180 evaluations to review (3 evaluations reviewed per day) would be most reasonable and that number is used in the LOE calculation below. However, the final list of reports to be reviewed will be submitted as part of the activity workplan.

### **3. Assessment Tools, Tool Orientation, and Inter-Rater Reliability Check**

The PPL/MSI tool has undergone several tests, including inter-rater reliability tests, and was improved as a result of these tests. The tool companion guide also helps raters to score evaluations in a more consistent manner. While we expect some modifications to the tool, we don't anticipate these will affect the tool reliability. However, to ensure raters have a good understanding of the tool and the scoring process, AFR will hold an orientation for the assessment team to go over the tool and answer any questions the team has. In addition, the team will be given an opportunity to pre-test the tool by applying it to one evaluation report not included in the sample to assess inter-rater reliability, identify

any major variations in rater scores, and revise or clarify the elements of the tool with inconsistent scores.

The Evaluation Report Quality Review checklist and rater's guide are publicly available in the USAID Meta-Evaluation report ([https://pdf.usaid.gov/pdf\\_docs/pdacx771.pdf](https://pdf.usaid.gov/pdf_docs/pdacx771.pdf)) and the E3 [Sectoral Synthesis of 2013 -2014 Evaluation Findings \(usaid.gov\)](#) (See Annex E, Page 158).

#### **4. Tool Administration**

It's anticipated that at least 180 evaluations will be reviewed. Each member of the assessment team will review, complete the checklist, and rate the quality of at least 60 evaluations (3 reports\* 60 days). AFR/SD staff who are available during the assessment period will be encouraged to participate in the evaluation review and rating of evaluations pertaining to their sectors. They will also add value to content analysis to identify lessons learned in project/activity design, implementation, and overall performance. If a reviewer has a question on one or more checklist criteria, he/she will be encouraged to communicate with other reviewers to see if they can clarify the criteria and if they are equally confused by the criteria, they should communicate with the Africa Bureau manager of this activity. In general, the right scores will be those provided by the majority of reviewers.

#### **5. Data Analysis**

Following completion of the checklist and scoring of reports, the assessment team will hold a mini-workshop with all the assessment team members to review the scores of each reviewer and discuss the data/information generated from content analysis or analysis of pre-determined report themes. This qualitative information should be captured during the report review stage in a format that is easy to analyze. The activity team will then systematically organize and analyze the data and prepare a draft report for AFR. AFR and the activity team will discuss and agree on the report template before the team drafts the report.

Statistical software may be used by the assessment team to analyze the data, including appropriate software for qualitative data. The synthesis report will include descriptive statistics on findings from the data set. The team will also conduct a trend analysis to determine the extent to which the quality of AFR evaluations have improved over time. Chi Square values may be used to assess correlation between the quality of evaluations with some of the key factors known to have a significant influence on evaluation quality.

### **B. Deliverables**

The following deliverables are envisioned as part of this activity. It's anticipated that the activity will have a 4-month period of performance, kick-off on/about July 15th and closing on/about November 15, 2022. A USDA partner implementing this activity will provide a detailed timeline and budget consistent with the above period of performance.

1. Draft and Final Activity Workplan and Budget
2. Updated Assessment Checklists and Content Analysis Questionnaire
3. Draft and Final Evaluation Quality and Synthesis Report incorporating programmatic lessons learned
4. Draft and Final Report Summary/Brief

## **VII. Team Composition**

It's anticipated that 3 evaluation specialists will be needed to complete the activities and deliverables listed above.

### ***Senior Evaluation Specialist/Team Leader (1)***

The Senior Evaluation Specialist/Team lead will oversee all the activities and production and delivery of activity deliverables. He will report to both KNUST and USDA managers of this project to ensure challenges related to this activity are addressed in a timely manner. He will ensure tasks are implemented as effectively and efficiently as possible. He will have the following qualifications:

- Master's degree in international development or related social science field
- At least 7 years of experience designing and conducting USAID or other major donor evaluations
- Experience as Evaluation Team lead preferred
- Strong knowledge/familiarity with USAID Evaluation Policy and evaluation quality standards
- Familiarity with USAID or other donor evaluation quality assessment tools
- General knowledge of USAID programming sectors including education, health, water and sanitation, democracy and governance, and economic growth (agriculture, climate change/environment, energy, and trade/investment)
- Excellent data synthesis, report writing, and communication/presentation skills

### ***Senior Evaluation Specialist (2)***

The 2 Senior Evaluation Specialists will work in tandem with the Activity Team lead to execute all the tasks for this assignment. They will have the following qualifications, similar to the Team lead's qualifications:

- Masters degree in international development or related social science field
- At least 7 years of experience designing and conducting USAID or other major donor evaluations
- Strong knowledge/familiarity with USAID Evaluation Policy and evaluation quality standards
- Familiarity with USAID or other donor evaluation quality assessment tools
- Experience analyzing quantitative and qualitative data using appropriate software
- General knowledge of USAID programming sectors including education, health, water and sanitation, democracy and governance, and economic growth (agriculture, climate change/environment, energy, and trade/investment)
- Excellent data synthesis, report writing, and communication/presentation skills

## ANNEX 2: EVALUATION REFERENCE LIST

The following table lists the 177 evaluations that were included in this assessment. Numbers are not sequential, as some evaluations were removed from the sample after the numbering system was established.

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
1	Ethiopia	Final performance evaluation of the Value Chain Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00XV33.pdf">https://pdf.usaid.gov/pdf_docs/PA00XV33.pdf</a>	2021	Ag and food security, Health	Performance
2	Ethiopia	Mid-term evaluation for Growth Through Nutrition	<a href="https://pdf.usaid.gov/pdf_docs/PA00XV31.pdf">https://pdf.usaid.gov/pdf_docs/PA00XV31.pdf</a>	2021	Ag and food security, Health, WASH	Performance
3	Madagascar	Mid-term evaluation of Rural Access to New Opportunities in Water, Sanitation, and Hygiene (RANOWASH) project	<a href="https://careevaluations.org/wp-content/uploads/RANO-WASH-Evaluation-Final-Report.pdf">https://careevaluations.org/wp-content/uploads/RANO-WASH-Evaluation-Final-Report.pdf</a>	2021	WASH	Performance
4	Tanzania	Data Driven Advocacy midterm performance evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00XVVF.pdf">https://pdf.usaid.gov/pdf_docs/PA00XVVF.pdf</a>	2021	DRG	Performance
5	Nigeria	Northern Education Initiatives Plus Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00XB17.pdf">https://pdf.usaid.gov/pdf_docs/PA00XB17.pdf</a>	2021	Education and Social Services	Performance
6	Malawi	LGAP Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00xm8v.pdf">https://pdf.usaid.gov/pdf_docs/PA00xm8v.pdf</a>	2021	DRG	Performance
7	Djibouti	Mid-term evaluation for "Djibouti Early Grade Reading Activity (DEGRA)"	<a href="https://pdf.usaid.gov/pdf_docs/PA00Z8H1.pdf">https://pdf.usaid.gov/pdf_docs/PA00Z8H1.pdf</a>	2021	Education and Social Services	Performance
8	Ethiopia	Evaluation for Transform Maternal and Child Health-focused Activities	<a href="https://pdf.usaid.gov/pdf_docs/PA00Z5BH.pdf">https://pdf.usaid.gov/pdf_docs/PA00Z5BH.pdf</a>	2021	Health	Impact
9	Ethiopia	End-line Evaluation Private Health Sector Program	<a href="https://pdf.usaid.gov/pdf_docs/PA00XDSF.pdf">https://pdf.usaid.gov/pdf_docs/PA00XDSF.pdf</a>	2021	Health	Performance
10	Ghana	Final Performance Evaluation of Agricultural Development and Value Chain Enhancement Project II (ADVANCE II)	<a href="https://pdf.usaid.gov/pdf_docs/PA00XDZF.pdf">https://pdf.usaid.gov/pdf_docs/PA00XDZF.pdf</a>	2021	Ag and food security	Performance
11	Mali	PERFORMANCE EVALUATION OF Girls Leadership and Empowerment through Education (GLEE)	<a href="https://pdf.usaid.gov/pdf_docs/PA00XR8J.pdf">https://pdf.usaid.gov/pdf_docs/PA00XR8J.pdf</a>	2021	Education and Social Services	Performance
12	Nigeria	E-WASH activity Midline Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00X8D7.pdf">https://pdf.usaid.gov/pdf_docs/PA00X8D7.pdf</a>	2021	WASH	Performance
13	Rwanda	Nguriza Nshore Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00Z2SP.pdf">https://pdf.usaid.gov/pdf_docs/PA00Z2SP.pdf</a>	2021	Ag and food security	Performance
14	Somalia	External Evaluation of Alternative Basic Education for Pastoralists and Other Out-of-School Children in Somalia activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00X9T5.pdf">https://pdf.usaid.gov/pdf_docs/PA00X9T5.pdf</a>	2021	Education and Social Services	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
15	Zambia	TGCC Zambia Climate Smart Agriculture Baseline Impact Evaluation and Tenure and Global Climate Change (TGCC) Endline Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00T79I.pdf">https://pdf.usaid.gov/pdf_docs/PA00T79I.pdf</a>	2021	Ag and food security	Impact
16	Washington	Performance Evaluation for the Power Africa Transactions and Reforms Program's (PATRP)	<a href="https://pdf.usaid.gov/pdf_docs/PA00WPFQ.pdf">https://pdf.usaid.gov/pdf_docs/PA00WPFQ.pdf</a>	2020	Energy and Infrastructure	Performance
17	AFR West Africa Regional	Voice for Peace	<a href="https://pdf.usaid.gov/pdf_docs/PA00W4G3.pdf">https://pdf.usaid.gov/pdf_docs/PA00W4G3.pdf</a>	2019	DRG	Impact
18	AFR West Africa Regional	WEST AFRICA FOOD SECURITY PARTNERSHIP (WAFSP) PROGRAM	<a href="https://pdf.usaid.gov/pdf_docs/PA00X2J4.pdf">https://pdf.usaid.gov/pdf_docs/PA00X2J4.pdf</a>	2018	Ag and food security	Performance
19	Ethiopia	End-term performance Evaluation for Youth Potential	<a href="https://pdf.usaid.gov/pdf_docs/PA00VJD2.pdf">https://pdf.usaid.gov/pdf_docs/PA00VJD2.pdf</a>	2020	Youth	Performance
20	Guinea	Performance Evaluation of the USAID Health Service Delivery	<a href="https://pdf.usaid.gov/pdf_docs/PA00XKTT.pdf">https://pdf.usaid.gov/pdf_docs/PA00XKTT.pdf</a>	2020	Health	Performance
21	Kenya	End Term Evaluation - Kenya Resilient Arid Lands Partnership for Integrated Development Program (Kenya RAPID)	<a href="https://pdf.usaid.gov/pdf_docs/PA00X8MC.pdf">https://pdf.usaid.gov/pdf_docs/PA00X8MC.pdf</a>	2020	Resilience, WASH	Performance
22	Malawi	Mid-term Evaluation of FTF Agricultural Diversification Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00X27V.pdf">https://pdf.usaid.gov/pdf_docs/PA00X27V.pdf</a>	2020	Ag and food security	Performance
23	Malawi	Fisheries Integration of Society and Habitats (FISH) Endline Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WGV4.pdf">https://pdf.usaid.gov/pdf_docs/PA00WGV4.pdf</a>	2020	Ag and food security	Performance
24	Malawi	Final Evaluation of Development Food Assistance Programs (DFAPs) -UBALE & NJIRA	<a href="https://pdf.usaid.gov/pdf_docs/PA00WRKR.pdf">https://pdf.usaid.gov/pdf_docs/PA00WRKR.pdf</a>	2020	Ag and food security	Performance
25	Malawi	Protecting Ecosystems and Restoring Forests in Malawi (PERFORM) Endline Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WKHD.pdf">https://pdf.usaid.gov/pdf_docs/PA00WKHD.pdf</a>	2020	NRM	Performance
26	Malawi	USAID/Malawi Country Development Cooperation Strategy (CDCS) Impact Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TNP3.pdf">https://pdf.usaid.gov/pdf_docs/PA00TNP3.pdf</a>	2020	N/A	Impact
27	Mali	EESA End line Evaluation Safe and equitable access to basic education (EESA)	<a href="https://pdf.usaid.gov/pdf_docs/PA00X2KJ.pdf">https://pdf.usaid.gov/pdf_docs/PA00X2KJ.pdf</a>	2020	Education and Social Services	Performance
28	Mali	Mid Term Evaluation of SIRA	<a href="https://pdf.usaid.gov/pdf_docs/PA00X2K9.pdf">https://pdf.usaid.gov/pdf_docs/PA00X2K9.pdf</a>	2020	Education and Social Services	Performance
29	Mali	End line evaluation of the Mali FTF Mali Cereal Value Chain and Livestock for Growth Programs	<a href="https://pdf.usaid.gov/pdf_docs/PA00WK8Q.pdf">https://pdf.usaid.gov/pdf_docs/PA00WK8Q.pdf</a>	2020	Ag and food security	Performance
30	Mozambique	Mid-term Performance Evaluation Feed the Future (FtF) Improved Seeds for Better Agriculture (SEMEAR)	<a href="https://pdf.usaid.gov/pdf_docs/PA00WHSC.pdf">https://pdf.usaid.gov/pdf_docs/PA00WHSC.pdf</a>	2020	Ag and food security	Performance
31	Mozambique	Feed the Future Mozambique Agricultural Innovations Activity (FTF INOVA)- Midterm Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00X4SD.pdf">https://pdf.usaid.gov/pdf_docs/PA00X4SD.pdf</a>	2020	Ag and food security	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
32	Nigeria	Electoral Empowerment for Civil Society Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00WM6V.pdf">https://pdf.usaid.gov/pdf_docs/PA00WM6V.pdf</a>	2020	DRG	Performance
33	Rwanda	Mid-Term Performance Evaluation of Feed the Future Rwanda Hinga Weze	<a href="https://pdf.usaid.gov/pdf_docs/PA00XDMN.pdf">https://pdf.usaid.gov/pdf_docs/PA00XDMN.pdf</a>	2020	Ag and food security	Performance
34	Rwanda	Mureke Dusome Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WCGJ.pdf">https://pdf.usaid.gov/pdf_docs/PA00WCGJ.pdf</a>	2020	Education and Social Services	Performance
35	Senegal	All Children Reading Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WRXZ.pdf">https://pdf.usaid.gov/pdf_docs/PA00WRXZ.pdf</a>	2020	Education and Social Services	Performance
36	Tanzania	Performance Evaluation of Public Sector Systems Strengthening Program (PS3)	<a href="https://www.measureevaluation.org/resources/publications/tre-19-26/at_download/document">https://www.measureevaluation.org/resources/publications/tre-19-26/at_download/document</a>	2020	DRG	Performance
37	Tanzania	Boesha Afya- Mid Term Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WG7M.pdf">https://pdf.usaid.gov/pdf_docs/PA00WG7M.pdf</a>	2020	Health	Performance
38	Uganda	Performance evaluation of the Feed the Future Youth Leadership for Agriculture (YLA) activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00WQWR.pdf">https://pdf.usaid.gov/pdf_docs/PA00WQWR.pdf</a>	2020	Ag and food security	Performance
39	Zimbabwe	Final Performance Evaluation of the ENSURE Development Food Assistance Project in Zimbabwe	<a href="https://pdf.usaid.gov/pdf_docs/PA00WMBI.pdf">https://pdf.usaid.gov/pdf_docs/PA00WMBI.pdf</a>	2020	Ag and food security	Performance
40	Zimbabwe	Final Performance Evaluation of the Amalima Development Food Assistance Project in Zimbabwe	<a href="https://pdf.usaid.gov/pdf_docs/PA00WHZ4.pdf">https://pdf.usaid.gov/pdf_docs/PA00WHZ4.pdf</a>	2020	Ag and food security	Performance
41	Rwanda	Mid-term Performance Evaluation: Soma Umenye Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00WNZX.pdf">https://pdf.usaid.gov/pdf_docs/PA00WNZX.pdf</a>	2019	Education and Social Services	Performance
42	AFR Southern Africa Regional	DRG Project Level Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W4NS.pdf">https://pdf.usaid.gov/pdf_docs/PA00W4NS.pdf</a>	2019	DRG	Performance
43	AFR Southern Africa Regional	Trade and Food Security Project Midterm Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W4N6.pdf">https://pdf.usaid.gov/pdf_docs/PA00W4N6.pdf</a>	2019	Ag and food security, Trade and Investment	Performance
44	AFR Southern Africa Regional	RESILIM Midterm Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W5H6.pdf">https://pdf.usaid.gov/pdf_docs/PA00W5H6.pdf</a>	2019	Environment and Climate Change, Resilience	Performance
45	AFR West Africa Regional	West Africa Biodiversity and Climate Change Program	<a href="https://pdf.usaid.gov/pdf_docs/PA00WF5R.pdf">https://pdf.usaid.gov/pdf_docs/PA00WF5R.pdf</a>	2019	Environment and climate change	Performance
46	AFR West Africa Regional	West African Health Informatics Team Sub-Activity Endline	<a href="https://pdf.usaid.gov/pdf_docs/PA00WCNI.pdf">https://pdf.usaid.gov/pdf_docs/PA00WCNI.pdf</a>	2020	Health	Performance
47	AFR West Africa Regional	West African Health Organization Leadership Capacity Strengthening (CAPS) Project Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W5H6.pdf">https://pdf.usaid.gov/pdf_docs/PA00W5H6.pdf</a>	2019	Health	Performance
48	AFR West Africa Regional	AfWA/WA-WASH	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJX4.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJX4.pdf</a>	2019	WASH	Performance
49	Ethiopia	Final performance evaluation of the Strengthening Ethiopia's Urban Health Program (SEUHP) activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJX4.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJX4.pdf</a>	2019	Health	Performance
50	Ethiopia	Final Performance Evaluation for Challenge TB	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJKQ.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJKQ.pdf</a>	2019	Health	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
51	Ghana	Impact Evaluation for Ghana Strengthening Accountability Mechanism Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJKS.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJKS.pdf</a>	2019	DRG	Impact
52	Ghana	Resiliency in Northern Ghana (RING)	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJKR.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJKR.pdf</a>	2019	Ag and food security, DRG, Health, WASH	Performance
53	Ghana	Final Performance Evaluation: Systems for Health Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJKQ.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJKQ.pdf</a>	2019	Health	Performance
54	Kenya	Mid-Term Evaluation - Strengthening Community Resilience Against Extremism (SCORE)	<a href="https://pdf.usaid.gov/pdf_docs/PA00TRDD.pdf">https://pdf.usaid.gov/pdf_docs/PA00TRDD.pdf</a>	2019	DRG	Performance
55	Kenya	Yetu Initiative - Mid-Term Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TQ59.pdf">https://pdf.usaid.gov/pdf_docs/PA00TQ59.pdf</a>	2019	DRG	Performance
56	Liberia	Democracy and Governance Whole-of-Project Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TVTC.pdf">https://pdf.usaid.gov/pdf_docs/PA00TVTC.pdf</a>	2019	DRG	Performance
57	Liberia	Partnership for Advancing Community Services (PACS) Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W8VD.pdf">https://pdf.usaid.gov/pdf_docs/PA00W8VD.pdf</a>	2019	WASH	Performance
58	Malawi	FTF - Malawi Strengthening Agricultural and Nutrition Extension Services (SANE) Mid-term Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W5VWV.pdf">https://pdf.usaid.gov/pdf_docs/PA00W5VWV.pdf</a>	2019	Ag and food security	Performance
59	Malawi	Impact Evaluation of the Wellness and Agriculture for Life Advancement (WALA) Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00W4GV.pdf">https://pdf.usaid.gov/pdf_docs/PA00W4GV.pdf</a>	2019	Ag and food security	Impact
60	Mozambique	Parceria para Boa Governacao Performance Evaluation PCBG	<a href="https://pdf.usaid.gov/pdf_docs/PA00TQ6T.pdf">https://pdf.usaid.gov/pdf_docs/PA00TQ6T.pdf</a>	2019	DRG	Performance
61	Mozambique	IFPP Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W5H5.pdf">https://pdf.usaid.gov/pdf_docs/PA00W5H5.pdf</a>	2019	Health	Performance
62	Niger	Participatory, Responsive Governance in Niger: Impact Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W6RS.pdf">https://pdf.usaid.gov/pdf_docs/PA00W6RS.pdf</a>	2019	DRG	Impact
63	Nigeria	Feed the Future Nigeria Livelihoods Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00W72N.pdf">https://pdf.usaid.gov/pdf_docs/PA00W72N.pdf</a>	2019	Ag and food security	Performance
64	Nigeria	FTF Nigeria Agricultural Policy project - midline evaluation 2018	<a href="https://pdf.usaid.gov/pdf_docs/PA00TWIX.pdf">https://pdf.usaid.gov/pdf_docs/PA00TWIX.pdf</a>	2019	Ag and food security	Performance
66	Rwanda	Impact Evaluation of Twiyubake program (former Improved Services for Vulnerable Populations (ISVP) Program)	<a href="https://pdf.usaid.gov/pdf_docs/PA00WC3S.pdf">https://pdf.usaid.gov/pdf_docs/PA00WC3S.pdf</a>	2019	Education and Social Services	Impact
67	Rwanda	Huguka Dukore Akazi Kanoze Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WGVG.pdf">https://pdf.usaid.gov/pdf_docs/PA00WGVG.pdf</a>	2019	Youth	Performance
68	Senegal	Impact Evaluation (final) of the agriculture and nutrition program (Yaajeende)	<a href="https://pdf.usaid.gov/pdf_docs/PA00TZ2D.pdf">https://pdf.usaid.gov/pdf_docs/PA00TZ2D.pdf</a>	2019	Ag and food security, Health	Impact
69	Senegal	Parliamentary Assistance & Civic Engagement/SUNU Budget Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TZ2B.pdf">https://pdf.usaid.gov/pdf_docs/PA00TZ2B.pdf</a>	2019	DRG	Performance
70	Senegal	ACCES WASH Mid-term Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TZ26.pdf">https://pdf.usaid.gov/pdf_docs/PA00TZ26.pdf</a>	2019	WASH	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
71	Sierra Leone	Performance Evaluation of the Women Empowered for Leadership and Development Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00W553.pdf">https://pdf.usaid.gov/pdf_docs/PA00W553.pdf</a>	2019	DRG	Performance
72	Somalia	TIS+ Baseline and Endline Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TW4W.pdf">https://pdf.usaid.gov/pdf_docs/PA00TW4W.pdf</a>	2019	DRG	Impact
74	Tanzania	Evaluation of the Pamoja Twajenga - Capacity Development for Partners of Accountability Program	<a href="https://pdf.usaid.gov/pdf_docs/PA00TT9R.pdf">https://pdf.usaid.gov/pdf_docs/PA00TT9R.pdf</a>	2019	DRG	Performance
75	Tanzania	Community Health And Social Welfare Program Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TRS2.pdf">https://pdf.usaid.gov/pdf_docs/PA00TRS2.pdf</a>	2019	Health	Performance
76	Tanzania	Tulonge Afya-Comprehensive Platform Integrated Communication Initiative Midterm Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WRZ7.pdf">https://pdf.usaid.gov/pdf_docs/PA00WRZ7.pdf</a>	2020	Health	Performance
77	Tanzania	SAUTI YETU Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TN9T.pdf">https://pdf.usaid.gov/pdf_docs/PA00TN9T.pdf</a>	2019	Health	Performance
78	Tanzania	Kizazi Kipya midterm evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TRS4.pdf">https://pdf.usaid.gov/pdf_docs/PA00TRS4.pdf</a>	2019	Education and Social Services	Performance
79	Uganda	Final Performance Evaluation of the Joint USAID/Uganda and DFID Uganda's Governance, Accountability, Participation and Performance Program (GAPPP)	<a href="https://pdf.usaid.gov/pdf_docs/PA00WB27.pdf">https://pdf.usaid.gov/pdf_docs/PA00WB27.pdf</a>	2019	DRG	Performance
80	Zambia	PROFIT+ EX POST EVALUATION	<a href="https://pdf.usaid.gov/pdf_docs/PA00WDR2.pdf">https://pdf.usaid.gov/pdf_docs/PA00WDR2.pdf</a>	2019	Ag and food security	Performance
81	AFR East Africa Regional	Mid-Term Evaluation of the Planning for Resilience in East Africa through Policy, Adaptation, Research and Economic Development (PREPARED)	<a href="https://pdf.usaid.gov/pdf_docs/PBAAJ790.pdf">https://pdf.usaid.gov/pdf_docs/PBAAJ790.pdf</a>	2017	Environment and climate change	Performance
82	AFR East Africa Regional	Final Evaluation of the Common Market for East and Southern Africa (COMESA)	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFHD.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFHD.pdf</a>	2018	Trade and Investment	Performance
83	AFR East Africa Regional	Mid-Term Evaluation of the East African Trade and Investment Hub (EATIH)	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFHF.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFHF.pdf</a>	2018	Trade and Investment	Performance
84	Uganda	Ex-Post Evaluation of USAID/Uganda Sustainable, Comprehensive Response for Orphans and Vulnerable Children and their Families (SCORE) Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00TJ6B.pdf">https://pdf.usaid.gov/pdf_docs/PA00TJ6B.pdf</a>	2018	Health, Resilience	Performance
85	Ghana	Ex-Post Evaluation of the Water Access, Sanitation and Hygiene for Urban Poor (WASH-UP) Activity in Ghana	<a href="https://pdf.usaid.gov/pdf_docs/PA00TJ6C.pdf">https://pdf.usaid.gov/pdf_docs/PA00TJ6C.pdf</a>	2018	WASH	Performance
86	AFR Sahel Regional	Final Evaluation of Sahel Resilience Learning (SAREL) Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00T7IH.pdf">https://pdf.usaid.gov/pdf_docs/PA00T7IH.pdf</a>	2018	Resilience	Performance
87	AFR West Africa Regional	Cotton Partnership (C4CP)	<a href="https://pdf.usaid.gov/pdf_docs/PA00T7TV.pdf">https://pdf.usaid.gov/pdf_docs/PA00T7TV.pdf</a>	2018	Ag and food security	Performance
88	AFR West Africa Regional	Pdev-II Endline Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00SWPK.pdf">https://pdf.usaid.gov/pdf_docs/PA00SWPK.pdf</a>	2018	DRG	Impact

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
89	Burkina Faso	Final evaluation of Victory against Malnutrition	<a href="https://pdf.usaid.gov/pdf_docs/PBAAJ791.pdf">https://pdf.usaid.gov/pdf_docs/PBAAJ791.pdf</a>	2018	Ag and food security	Performance
90	Ethiopia	End-term performance evaluation of READ TA	<a href="https://pdf.usaid.gov/pdf_docs/PA00SZB8.pdf">https://pdf.usaid.gov/pdf_docs/PA00SZB8.pdf</a>	2018	Education and Social Services	Performance
91	Ethiopia	Final performance evaluation for strengthening human resource for health	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDV1.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDV1.pdf</a>	2018	Health	Performance
92	Kenya	Kenya Agriculture Value Chains Enterprises Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00T9H6.pdf">https://pdf.usaid.gov/pdf_docs/PA00T9H6.pdf</a>	2018	Ag and food security	Performance
93	Kenya	Agile and Harmonized Assistance for Devolved Institutions (AHADI) Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFHN.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFHN.pdf</a>	2018	DRG	Performance
94	Kenya	Kenya Innovation Engine Final Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDMF.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDMF.pdf</a>	2018	Ag and food security	Performance
95	Kenya	End of Project Evaluation of APHIAplus Imarisha	<a href="https://pdf.usaid.gov/pdf_docs/PA00TMG2.pdf">https://pdf.usaid.gov/pdf_docs/PA00TMG2.pdf</a>	2018	Health	Performance
96	Liberia	Midterm Evaluation of Land Governance Support Activity (LGSA)	<a href="https://pdf.usaid.gov/pdf_docs/PA00T5GQ.pdf">https://pdf.usaid.gov/pdf_docs/PA00T5GQ.pdf</a>	2018	DRG	Performance
97	Liberia	Legal Professional Development and Anti-Corruption (LPAC) Midterm Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00THDS.pdf">https://pdf.usaid.gov/pdf_docs/PA00THDS.pdf</a>	2018	DRG	Performance
98	Liberia	Liberia Municipal Water Project II (LMWP II) Midterm and Capital Improvement Project (CIP) Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W5B9.pdf">https://pdf.usaid.gov/pdf_docs/PA00W5B9.pdf</a>	2018	Infrastructure	Performance
99	Malawi	Malawi Early Grade Reading Project Impact Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00T3Q6.pdf">https://pdf.usaid.gov/pdf_docs/PA00T3Q6.pdf</a>	2018	Education and Social Services	Impact
100	Malawi	ASPIRE Mid-Term Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00SYSB.pdf">https://pdf.usaid.gov/pdf_docs/PA00SYSB.pdf</a>	2018	Education and Social Services, Health	Performance
101	Mali	Performance Evaluation of the IRP for Nutrition Program	<a href="https://pdf.usaid.gov/pdf_docs/PA00SXW4.pdf">https://pdf.usaid.gov/pdf_docs/PA00SXW4.pdf</a>	2018	Ag and food security, WASH	Performance
102	Mozambique	Local Capacity development Activity Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TG3H.pdf">https://pdf.usaid.gov/pdf_docs/PA00TG3H.pdf</a>	2018	DRG	Performance
103	Mozambique	FINAL PERFORMANCE EVALUATION USAID/MOZAMBIQUE SOCIAL AND BEHAVIOR CHANGE COMMUNICATION PROGRAM (PIRCOM)	<a href="https://pdf.usaid.gov/pdf_docs/PA00TJM9.pdf">https://pdf.usaid.gov/pdf_docs/PA00TJM9.pdf</a>	2018	Health	Performance
104	Nigeria	Maximizing Agricultural Revenues and Key Enterprises in Targeted Sites (MARKETS) II	<a href="https://pdf.usaid.gov/pdf_docs/PA00T6FM.pdf">https://pdf.usaid.gov/pdf_docs/PA00T6FM.pdf</a>	2018	Ag and food security	Performance
105	Nigeria	Nigeria Renewable Energy and Energy Efficiency Project (REEEP) Final Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDS5.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDS5.pdf</a>	2018	Energy and Infrastructure	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
106	Nigeria	Promoting the Quality of Medicines (PQM) Midterm Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDS6.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDS6.pdf</a>	2018	Health	Performance
107	Rwanda	Mid-Term Performance Evaluation for Private Sector Driven Agricultural Growth (PSD-AG)	<a href="https://pdf.usaid.gov/pdf_docs/PA00T981.pdf">https://pdf.usaid.gov/pdf_docs/PA00T981.pdf</a>	2018	Ag and food security	Performance
108	Rwanda	SOMA UMENYE Impact Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00T13Z.pdf">https://pdf.usaid.gov/pdf_docs/PA00T13Z.pdf</a>	2018	Education and Social Services	Impact
109	Senegal	Women's Leadership/Civic Journalism Activity Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDSZ.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDSZ.pdf</a>	2017	DRG	Performance
110	Senegal	G2G PASMESISS Project-HSS related Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFQH.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFQH.pdf</a>	2018	Health	Performance
111	South Africa	STELLAR/ELIT Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00T61D.pdf">https://pdf.usaid.gov/pdf_docs/PA00T61D.pdf</a>	2018	Education and Social Services	Impact
112	Tanzania	Rule of Law: Access to Justice Integrating Gender-Based Violence	<a href="https://pdf.usaid.gov/pdf_docs/PA00T7RM.pdf">https://pdf.usaid.gov/pdf_docs/PA00T7RM.pdf</a>	2018	DRG	Performance
113	Tanzania	TUSOME PAMOJA performance evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TGR7.pdf">https://pdf.usaid.gov/pdf_docs/PA00TGR7.pdf</a>	2018	Education and Social Services	Performance
114	Tanzania	Tanzania Rural Roads Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TF6R.pdf">https://pdf.usaid.gov/pdf_docs/PA00TF6R.pdf</a>	2018	Energy and Infrastructure	Performance
115	Tanzania	WARIDI WASH Midterm Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFD7.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFD7.pdf</a>	2018	WASH	Performance
116	Uganda	Performance Evaluation of the Supporting Access to Justice, Fostering Peace and Equity	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDFC.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDFC.pdf</a>	2018	DRG	Performance
117	Uganda	Evaluation of the USAID Uganda Private Health Support (PHS) Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDFC.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDFC.pdf</a>	2018	Health	Performance
118	Uganda	Evaluation of the Communication for Healthy Communities (CHC) Activity	<a href="https://pdf.usaid.gov/pdf_docs/PA00TDP5.pdf">https://pdf.usaid.gov/pdf_docs/PA00TDP5.pdf</a>	2018	Health	Performance
119	Zambia	Gender and Groundnut Value Chain Impact Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TCNM.pdf">https://pdf.usaid.gov/pdf_docs/PA00TCNM.pdf</a>	2018	Ag and food security	Impact
120	Zambia	Community Forest project Performance Evaluation.	<a href="https://pdf.usaid.gov/pdf_docs/PA00TJ62.pdf">https://pdf.usaid.gov/pdf_docs/PA00TJ62.pdf</a>	2018	NRM	Performance
121	AFR Central Africa Region	Mid-term Evaluation of the Third Phase of Central Africa Regional Program for the Environment (CARPE III)	<a href="https://pdf.usaid.gov/pdf_docs/PA00MZ8M.pdf">https://pdf.usaid.gov/pdf_docs/PA00MZ8M.pdf</a>	2017	Environment and Climate Change, NRM	Performance
123	AFR Southern Africa Regional	Evaluation of the Beyond Advocacy Fund (BAF)	<a href="https://pdf.usaid.gov/pdf_docs/pa00mz9t.pdf">https://pdf.usaid.gov/pdf_docs/pa00mz9t.pdf</a>	2017	Education and Social Services	Performance
124	AFR West Africa Regional	Mid-term Performance Evaluation of the West Africa Trade Hub and African Partners Network	<a href="https://pdf.usaid.gov/pdf_docs/PA00N494.pdf">https://pdf.usaid.gov/pdf_docs/PA00N494.pdf</a>	2017	Trade and investment	Performance
125	AFR West Africa Regional	Mid-term Performance Evaluation of Agir pour la Planification Familiale (AGIR PF)-- Regional Family Planning Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00SS59.pdf">https://pdf.usaid.gov/pdf_docs/PA00SS59.pdf</a>	2017	Health	Performance
126	AFR West Africa Regional	SCMS Final Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N35G.pdf">https://pdf.usaid.gov/pdf_docs/PA00N35G.pdf</a>	2017	Health	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
127	AFR West Africa Regional	Midterm Evaluation of the Sanitation Service Delivery (SSD) Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00N4Z5.pdf">https://pdf.usaid.gov/pdf_docs/PA00N4Z5.pdf</a>	2017	WASH	Performance
128	Angola	Performance Evaluation of the Systems for Improved Access to Pharmaceuticals and Services (SIAPS)	<a href="https://pdf.usaid.gov/pdf_docs/PA00SVK8.pdf">https://pdf.usaid.gov/pdf_docs/PA00SVK8.pdf</a>	2017	Health	Performance
129	Kenya	Tusome (Early Grade Reading) Program Endline Performance Evaluation	<a href="https://www.norc.org/PDFs/Tusome%20Endline%20Performance%20Evaluation/PA00XVBP.pdf">https://www.norc.org/PDFs/Tusome%20Endline%20Performance%20Evaluation/PA00XVBP.pdf</a>	2021	Education and Social Services	Performance
130	Kenya	Mid term Performance Evaluation - Kenya Integrated Water Sanitation and Hygiene (KIWASH) Activity	<a href="https://www.globalwaters.org/sites/default/files/kiwash-midterm-evaluation.pdf">https://www.globalwaters.org/sites/default/files/kiwash-midterm-evaluation.pdf</a>	2019	WASH	Performance
131	Nigeria	Strengthening Advocacy and Civic Engagement (SACE) Follow-Up Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00W6Z4.pdf">https://pdf.usaid.gov/pdf_docs/PA00W6Z4.pdf</a>	2019	DRG	Performance
132	Nigeria	E2A-SMGL End of Project Evaluation	<a href="https://culturalpractice.com/wp-content/uploads/Nigeria-Saving-Mothers-Giving-Life-SMGL-End-of-Project-Evaluation.pdf">https://culturalpractice.com/wp-content/uploads/Nigeria-Saving-Mothers-Giving-Life-SMGL-End-of-Project-Evaluation.pdf</a>	2019	Health	Performance
133	Senegal	Mid-Term Performance Evaluation of Health Project	<a href="https://pdf.usaid.gov/pdf_docs/PA00W8Z5.pdf">https://pdf.usaid.gov/pdf_docs/PA00W8Z5.pdf</a>	2019	Health	Performance
134	Senegal	Governance for Local Development mid-term performance evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00WJNF.pdf">https://pdf.usaid.gov/pdf_docs/PA00WJNF.pdf</a>	2019	DRG, Health, WASH	Performance
135	Uganda	Malaria Action Program for Districts (MAPD) Midterm Learning Review	<a href="https://pdf.usaid.gov/pdf_docs/PA00W7TP.pdf">https://pdf.usaid.gov/pdf_docs/PA00W7TP.pdf</a>	2019	Health	Performance
136	Niger	Niger Education and Community Strengthening Project: Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFQN.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFQN.pdf</a>	2018	DRG, Education and Social Services	Performance
137	Mali	Evaluation of the High Impact Health Services (SSGI) and the Social Marketing/Behavior Change Communication (KJK) Programs	<a href="https://pdf.usaid.gov/pdf_docs/PA00THNT.pdf">https://pdf.usaid.gov/pdf_docs/PA00THNT.pdf</a>	2018	Health	Performance
138	Tanzania	EgenderHealth RESPOND performance evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00T7RM.pdf">https://pdf.usaid.gov/pdf_docs/PA00T7RM.pdf</a>	2018	Health	Performance
139	Ghana	Performance Evaluation: Accountable Democratic Institutions System strengthening (ADISS)	<a href="https://pdf.usaid.gov/pdf_docs/PA00N24F.pdf">https://pdf.usaid.gov/pdf_docs/PA00N24F.pdf</a>	2018	DRG	Performance
140	Burkina Faso	Final Evaluation of Families Achieving Sustainable Outcomes (FASO) program	<a href="https://pdf.usaid.gov/pdf_docs/PA00N51N.pdf">https://pdf.usaid.gov/pdf_docs/PA00N51N.pdf</a>	2017	Ag and food security, Health	Performance
141	Burkina Faso	Performance evaluation of Partnership for Participation and Poise program	<a href="https://pdf.usaid.gov/pdf_docs/PA00N5S8.pdf">https://pdf.usaid.gov/pdf_docs/PA00N5S8.pdf</a>	2017	DRG	Performance
142	Burundi	Mid-Term Evaluation of the AMASHIGA USAID\FFP Title II Program	<a href="https://pdf.usaid.gov/pdf_docs/PA00TD9R.pdf">https://pdf.usaid.gov/pdf_docs/PA00TD9R.pdf</a>	2017	Ag and food security	Performance
143	Cote d'Ivoire	Evaluation of the Projustice Program	<a href="https://pdf.usaid.gov/pdf_docs/pa00mt44.pdf">https://pdf.usaid.gov/pdf_docs/pa00mt44.pdf</a>	2017	DRG	Performance
144	Ethiopia	Mid term evaluation of Enhancing the Status of Human Rights Protection and Systems of Good Governance in Ethiopia	<a href="https://pdf.usaid.gov/pdf_docs/pa00n4z4.pdf">https://pdf.usaid.gov/pdf_docs/pa00n4z4.pdf</a>	2017	DRG	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
145	Ethiopia	Mid-term Performance Evaluation of the Health Sector Finance Reform (HSFR)	<a href="https://pdf.usaid.gov/pdf_docs/pa00n526.pdf">https://pdf.usaid.gov/pdf_docs/pa00n526.pdf</a>	2017	Health	Performance
146	Ethiopia	Final Performance Evaluation of Graduation with Resilience to Achieve Sustainable Development (GRAD) Project	<a href="https://pdf.usaid.gov/pdf_docs/pa00mxkk.pdf">https://pdf.usaid.gov/pdf_docs/pa00mxkk.pdf</a>	2017	Resilience	Performance
147	Ethiopia	Final Performance evaluation of the Community-based Disaster Risk Reduction activities (REVIVE and REAAP)	<a href="https://pdf.usaid.gov/pdf_docs/PA00MXKH.pdf">https://pdf.usaid.gov/pdf_docs/PA00MXKH.pdf</a>	2017	Environment and Climate Change, Resilience	Performance
148	Ethiopia	Final performance evaluation for the Yekokeb Berhan Program for Highly Vulnerable Children	<a href="https://pdf.usaid.gov/pdf_docs/pa00n63m.pdf">https://pdf.usaid.gov/pdf_docs/pa00n63m.pdf</a>	2017	Education and Social Services	Performance
149	Ghana	Partnership for Growth - Ghana Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/pa00st75.pdf">https://pdf.usaid.gov/pdf_docs/pa00st75.pdf</a>	2017	Energy and Infrastructure	Performance
150	Kenya	Final Performance Evaluation of Northern Rangelands Trust (NRT)	<a href="https://pdf.usaid.gov/pdf_docs/pa00mng5.pdf">https://pdf.usaid.gov/pdf_docs/pa00mng5.pdf</a>	2017	NRM	Performance
151	Kenya	Tusome (Early Grade Reading) Activity Midline Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00MS6J.pdf">https://pdf.usaid.gov/pdf_docs/PA00MS6J.pdf</a>	2017	Education and Social Services	Performance
152	Liberia	Smallholder Oil-Palm Support Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N772.pdf">https://pdf.usaid.gov/pdf_docs/PA00N772.pdf</a>	2017	Ag and food security	Performance
153	Liberia	Final Evaluation of Excellence in Higher Education for Liberian Development	<a href="https://pdf.usaid.gov/pdf_docs/pa00mwXH.pdf">https://pdf.usaid.gov/pdf_docs/pa00mwXH.pdf</a>	2017	Education and Social Services	Performance
154	Liberia	Education Quality and Access in Liberia Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/pa00mzz7.pdf">https://pdf.usaid.gov/pdf_docs/pa00mzz7.pdf</a>	2017	Education and Social Services	Performance
155	Liberia	Mid term Evaluation of Collaborative Support for Health (CSH)	<a href="https://pdf.usaid.gov/pdf_docs/PA00MXTW.pdf">https://pdf.usaid.gov/pdf_docs/PA00MXTW.pdf</a>	2017	Health	Performance
156	Liberia	Fixed Amount Reimbursement Agreement Final Evaluation Health Outputs Study	<a href="https://pdf.usaid.gov/pdf_docs/pa00mx61.pdf">https://pdf.usaid.gov/pdf_docs/pa00mx61.pdf</a>	2017	Health	Performance
157	Liberia	Fixed Amount Reimbursement Agreement Final Evaluation Capacity Assessment	<a href="https://pdf.usaid.gov/pdf_docs/PA00MPJT.pdf">https://pdf.usaid.gov/pdf_docs/PA00MPJT.pdf</a>	2017	Health	Performance
158	Liberia	Mid-term Evaluation of PACS	<a href="https://pdf.usaid.gov/pdf_docs/pa00n2f1.pdf">https://pdf.usaid.gov/pdf_docs/pa00n2f1.pdf</a>	2017	Health	Performance
159	Madagascar	Food for Peace Madagascar Development Portfolio Mid-term Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/pa00svqd.pdf">https://pdf.usaid.gov/pdf_docs/pa00svqd.pdf</a>	2017	Ag and food security	Performance
160	Madagascar	Integrated Social Marketing Performance Final Evaluation	<a href="https://marketbookshelf.com/publications/final-performance-evaluation-usaidmadagascar-integrated-social-marketing-program/">https://marketbookshelf.com/publications/final-performance-evaluation-usaidmadagascar-integrated-social-marketing-program/</a>	2017	Health	Performance
161	Malawi	Protecting Ecosystems and Restoring Forests in Malawi (PERFORM) Performance Evaluation-Mid-term	<a href="https://pdf.usaid.gov/pdf_docs/pa00n4jn.pdf">https://pdf.usaid.gov/pdf_docs/pa00n4jn.pdf</a>	2017	NRM	Performance
162	Malawi	End of project evaluation of the "Community Outreach Family Planning project"	<a href="https://pdf.usaid.gov/pdf_docs/pa00n3zb.pdf">https://pdf.usaid.gov/pdf_docs/pa00n3zb.pdf</a>	2017	Health	Performance

#	Commissioner	Full Title	Evaluation URL	Year	Commissioning Sector(s)	Evaluation Type
163	Malawi	Fisheries Integration of Society and Habitats (FISH) Performance Evaluation (Mid-Term)	<a href="https://pdf.usaid.gov/pdf_docs/pa00n589.pdf">https://pdf.usaid.gov/pdf_docs/pa00n589.pdf</a>	2017	Environment and climate change	Performance
164	Mali	Mid-Term Evaluation of the Cereal Value Chain (CVC) Project	<a href="https://pdf.usaid.gov/pdf_docs/pa00mp4q.pdf">https://pdf.usaid.gov/pdf_docs/pa00mp4q.pdf</a>	2017	Ag and food security	Performance
165	Mozambique	Coastal City Adaptation Project Mid-term Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/pa00mjg4.pdf">https://pdf.usaid.gov/pdf_docs/pa00mjg4.pdf</a>	2017	Resilience	Performance
166	Nigeria	Midterm Evaluation: Strengthening Advocacy and Civic Engagement (SACE)	<a href="https://pdf.usaid.gov/pdf_docs/pa00n6qf.pdf">https://pdf.usaid.gov/pdf_docs/pa00n6qf.pdf</a>	2017	DRG	Performance
167	Nigeria	Nigeria Expanded Trade and Transport (NEXTT) Project Final Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/pa00mjg4.pdf">https://pdf.usaid.gov/pdf_docs/pa00mjg4.pdf</a>	2017	Trade and investment	Performance
168	Nigeria	Northern Education Initiative Plus Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N894.pdf">https://pdf.usaid.gov/pdf_docs/PA00N894.pdf</a>	2017	Education and Social Services	Performance
169	Senegal	Final Performance Evaluation of the SCOPE and DIRECT Activities in Senegal	<a href="https://pdf.usaid.gov/pdf_docs/pa00n5ss.pdf">https://pdf.usaid.gov/pdf_docs/pa00n5ss.pdf</a>	2017	DRG	Performance
170	South Africa	External Evaluation of the USAID/South Africa "Systems Strengthening for Better HIV/TB Outcomes" Project (2012-2017) Comprehensive HIV District Support Program.	<a href="https://pdf.usaid.gov/pdf_docs/PA00SVTK.pdf">https://pdf.usaid.gov/pdf_docs/PA00SVTK.pdf</a>	2017	Health	Performance
171	South Sudan	FINAL PERFORMANCE EVALUATION BOMA-JONGLEI-EQUATORIA LANDSCAPE (BJEL) 2017 PROGRAM	<a href="https://pdf.usaid.gov/pdf_docs/PA00SYGV.pdf">https://pdf.usaid.gov/pdf_docs/PA00SYGV.pdf</a>	2017	NRM	Performance
172	South Sudan	Abyei Rehabilitation Initiative Phase II	<a href="https://pdf.usaid.gov/pdf_docs/PA00N6Q7.pdf">https://pdf.usaid.gov/pdf_docs/PA00N6Q7.pdf</a>	2017	DRG	Performance
173	Tanzania	TIBU HOMA performance evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00MR5Q.pdf">https://pdf.usaid.gov/pdf_docs/PA00MR5Q.pdf</a>	2017	Health	Performance
174	Tanzania	Natural Resource Management Performance Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N4SR.pdf">https://pdf.usaid.gov/pdf_docs/PA00N4SR.pdf</a>	2017	NRM	Performance
175	Uganda	USAID/Uganda's Systems for Treating AIDS Nationally (SUSTAIN) project	<a href="https://pdf.usaid.gov/pdf_docs/pa00mkms.pdf">https://pdf.usaid.gov/pdf_docs/pa00mkms.pdf</a>	2017	Health	Performance
177	Zimbabwe	MCHIP Associate Award Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N6B5.pdf">https://pdf.usaid.gov/pdf_docs/PA00N6B5.pdf</a>	2017	Health	Performance
178	Zimbabwe	Families and Communities for Elimination (FACE) of Pediatric HIV Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00N4XX.pdf">https://pdf.usaid.gov/pdf_docs/PA00N4XX.pdf</a>	2017	Health	Performance
179	Zimbabwe	PERFORMANCE EVALUATION OF USAID/ZIMBABWE'S IMPROVING FAMILY PLANNING SERVICES (IFPS) ACTIVITY	<a href="https://pdf.usaid.gov/pdf_docs/PA00N6B4.pdf">https://pdf.usaid.gov/pdf_docs/PA00N6B4.pdf</a>	2017	Health	Performance
180	Senegal	Fixed-Amount Reimbursement Agreement (FARA): National Malaria Control Program (NMCP) Evaluation	<a href="https://pdf.usaid.gov/pdf_docs/PA00TFQM.pdf">https://pdf.usaid.gov/pdf_docs/PA00TFQM.pdf</a>	2017	Health	Performance
181	AFR East Africa Regional	Final Evaluation - Cross-Border Health Integrated Partnership (CB-HIPP)	<a href="#">_____</a>	2019	Health	Performance

## ANNEX 3: LIST OF DOCUMENTS REFERENCED

- Alvares de Azevdo, Thomaz. 2018. *Assessment of the Quality of USAID-Funded Evaluations: Education Sector, 2013-2016*. [https://pdf.usaid.gov/pdf\\_docs/pa00srw1.pdf](https://pdf.usaid.gov/pdf_docs/pa00srw1.pdf)
- Freudenberer, Elizabeth. 2015. *Sectoral Synthesis of 2013-2014 Evaluation Findings: Bureau for Economic Growth, Education, and Environment*. [https://www.usaid.gov/sites/default/files/documents/1865/E3\\_Sectoral\\_Synthesis\\_Report.pdf](https://www.usaid.gov/sites/default/files/documents/1865/E3_Sectoral_Synthesis_Report.pdf)
- Green, Andrew, and Sam Hargadine. 2017. *Sectoral Synthesis of FY2015 Evaluation Findings: Bureau for Economic Growth, Education, and Environment*. [https://pdf.usaid.gov/pdf\\_docs/PA00MPI7.pdf](https://pdf.usaid.gov/pdf_docs/PA00MPI7.pdf)
- Hageboeck, Molly, Micah Frumkin, and Stephanie Monchein. 2013. *Meta-Evaluation of Quality and Coverage of USAID Evaluations, 2019-2021*. <https://www.usaid.gov/evaluation/meta-evaluation-quality-and-coverage>
- Hageboeck, Molly, Elizabeth Freudenberger, and Jason Catanach. 2019. *Learning from Evaluation Syntheses: Technical Report on Making Evidence From Evaluations More Accessible To Decision-Makers*. [https://usaidlearninglab.org/sites/default/files/resource/files/technical\\_report\\_learning\\_from\\_evaluation\\_syntheses.pdf](https://usaidlearninglab.org/sites/default/files/resource/files/technical_report_learning_from_evaluation_syntheses.pdf)
- Synthesis of Evaluations Related to the Feed the Future Learning Agenda*. 2016. <https://agrilinks.org/sites/default/files/resource/files/Final%20KDDAD%20Evaluation%20Synthesis.pdf>
- USAID Bureau for Policy, Planning and Learning, Office of Learning, Evaluation and Research. 2011. *USAID Evaluation Policy: January 2011*. <https://www.usaid.gov/sites/default/files/documents/2151/USAIDEvaluationPolicy.pdf>
- USAID Bureau for Policy, Planning and Learning, Office of Learning, Evaluation and Research. 2016. *USAID Evaluation Policy: January 2011, Updated October 2016*. <https://2017-2020.usaid.gov/sites/default/files/documents/1870/USAIDEvaluationPolicy.pdf>
- USAID Bureau for Policy, Planning and Learning, Office of Learning, Evaluation and Research. 2020. *USAID Evaluation Policy: October 2020*. [https://www.usaid.gov/sites/default/files/documents/Evaluation\\_Policy\\_Update\\_OCT2020\\_Final.pdf](https://www.usaid.gov/sites/default/files/documents/Evaluation_Policy_Update_OCT2020_Final.pdf)
- USAID Bureau for Policy, Planning and Learning. 2013. *How-To Note: Evaluation Statements of Work*. [https://www.usaid.gov/sites/default/files/documents/1865/508\\_Howto\\_Note\\_Evaluation\\_Statements\\_Work.pdf](https://www.usaid.gov/sites/default/files/documents/1865/508_Howto_Note_Evaluation_Statements_Work.pdf)
- USAID Bureau for Policy, Planning and Learning. 2021. *Discussion Note: Making Evidence Accessible through Evaluation Synthesis Platforms*. [https://usaidlearninglab.org/sites/default/files/resource/files/dn-evidence\\_eval\\_synthesis\\_final2021\\_1.pdf](https://usaidlearninglab.org/sites/default/files/resource/files/dn-evidence_eval_synthesis_final2021_1.pdf)
- USAID 2016. *USAID Evaluation Report Requirements: A Mandatory Reference for ADS Chapter 201*. <https://www.usaid.gov/sites/default/files/documents/1868/201mah.pdf>
- USAID. 2022. *Criteria to Ensure the Quality of the Evaluation Report: A Mandatory Reference for ADS Chapter 201*. <https://www.usaid.gov/sites/default/files/documents/201maa.pdf>

## ANNEX 4: DATA COLLECTION TOOLS

### Evaluation Quality Checklist

Question	Yes	No	N/A <sup>37</sup>
<b>Executive summary</b>			
1. Does the Executive Summary accurately reflect the most critical elements of the report?			
<b>Program/Project Background</b>			
2. Are the basic characteristics of the program, project or activity described (title, dates, funding organization, budget, implementing organization, location/map, target group, contextual information)?			
3. Is the program or project's "theory of change" described (intended results (in particular the project purpose); development hypotheses; assumptions)			
<b>Evaluation Purpose</b>			
4. Does the evaluation purpose identify the management reason(s) for undertaking the evaluation?			
<b>Evaluation Questions</b>			
5. Are the evaluation questions stated in the body of the report clearly related to the evaluation purpose?			
6. Are the evaluation questions in the report identical to the evaluation questions in the evaluation SOW?			
7. If the questions in the body of the report and those found in the SOW differ, does the report (or annexes) state that there was written approval for changes in the evaluation questions?			
<b>Methodology</b>			
8. Does the report (or methods annex) describe specific data collection methods the team used?			
9. Are the data collection methods presented (in the report or methods annex) in a manner that makes it clear which specific methods are used to address each evaluation question? (e.g., matrix of questions by methods)			
10. Does the report (or methods annex) describe specific data analysis methods the team used? (frequency distributions, cross-tabulations; correlation; reanalysis of secondary data)			

<sup>37</sup> N/A = conditions required to answer the question are not all present.

Question	Yes	No	N/A <sup>37</sup>
11. Are the data analysis methods presented (in the report or methods annex) in a manner that makes it clear how they are associated with the evaluation questions or specific data collection methods?			
<b>Team Composition</b>			
12. Did the report (or methods annex) indicate that the evaluation team leader was external to USAID?			
13. Did the report (or methods annex) identify at least one evaluation specialist on the team?			
14. Did the report (or methods annex) identify local evaluation team members?			
15. Did the report indicate that team members had signed Conflict of Interest forms or letters? (check if the report says this or the COI forms are included in an annex)			
<b>Study Limitations</b>			
16. Does the report include a description of study limitations (lack of baseline data; selection bias as to sites, interviewees, comparison groups; seasonal unavailability of key informants)?			
<b>Responsiveness to Evaluation Questions</b>			
17. Is the evaluation report structured to present findings in relation to evaluation questions, as opposed to presenting information in relation to program/project objectives or in some other format?			
18. Are all of the evaluation questions, including sub-questions, answered primarily in the body of the report (as opposed to in an annex)			
19. If any questions were not answered, did the report provide a reason why?			
<b>Findings</b>			
20. Did the findings presented appear to be drawn from social science data collection and analysis methods the team described in its study methodology (including secondary data it assembled or reanalyzed)?			
21. For findings presented within the evaluation report is there a transparent connection to the source(s) of the data? (60% of the beneficiaries' interviews reported that...)			
22. In the presentation of findings, did the team draw on data from the range of methods they used rather than answer using data from primarily one method?			
23. Are findings clearly distinguished from conclusions and recommendations in the report, at least by the use of language that signals transitions (“the evaluation found that....” “the team concluded that ....”)?			

Question	Yes	No	N/A <sup>37</sup>
24. Are quantitative findings reported precisely, i.e., as specific numbers or percentages rather than general statements like “some”, “many”, or “most”?			
25. Does the report present findings about unplanned/unanticipated results?			
26. Does the report discuss alternative possible causes of results/outcomes it documents?			
27. Are evaluation findings disaggregated by sex at all levels (activity, outputs, outcomes) when data are person-focused?			
28. Does the report explain whether access/participation and/or outcomes/benefits were different for men and women when data are person-focused?			
<b>Recommendations</b>			
29. Is the report’s presentation of recommendations limited to recommendations? (free from repetition of information already presented or new findings not previously revealed)			
30. Do evaluation recommendations meet USAID policy expectations with respect to being specific? (states clearly what is to be done, and possibly how?)			
31. Do evaluation recommendations meet USAID policy expectations with respect to being directed to a specific party? (identifies who should do it)			
32. Are all the recommendations supported by the findings and conclusions presented? (Can a reader can follow a transparent path from findings to conclusions to recommendations?)			
<b>Annexes</b>			
33. Is the evaluation SOW included as an annex to the evaluation report?			
34. Are sources of information that the evaluators used listed in annexes?			
35. Are data collection instruments provided as evaluation report annexes?			
36. Is there a matching instrument for each and every data collection method the team reported that they used?			
37. Were any “Statements of Differences” included as evaluation annexes (prepared by team members, the Mission, the Implementing Partner, or other stakeholder)?			
<b>Evaluation Data Warehousing</b>			
38. Does the evaluation report explain how/in what form the evaluation data will be transferred to USAID (survey data, focus group transcripts)?			
<b>Link to Evaluation Policy Quality Standards</b>			

Question	Yes	No	N/A <sup>37</sup>
39. Does the evaluation SOW include a copy or the equivalent of Appendix I of the evaluation policy?			
<b>Additional questions about basic evaluation characteristics</b>			
40. Does the report include a Table of Contents?			
41. Does the report include a glossary and/or list of acronyms?			
42. Is the report well-written (clear sentences, reasonable length paragraphs) and mostly free of typos and other grammatical errors?			
43. Is the report well-organized (each topic is clearly delineated, subheadings used for easy reading)?			
44. Is the date of the report given on the report cover or inside cover?			
45. Is the name of the team leader present in the report or on the report cover, inside cover, or in the preface or introduction to the report?			

## Evaluation Quality Checklist Raters' Guide

<b>Executive Summary</b>	
1. Does the executive summary present an accurate reflection of the most critical elements of the report?	An executive summary must provide an accurate representation of the gist of the evaluation report without adding any new “material” information or contradicting the evaluation report in any way. “Critical” implies that not all information included in the evaluation report needs to be present in the executive summary, but that critical information from <u>all major elements should be discussed (i.e., evaluation purpose, questions, background information, methods, study limitations, findings, and recommendations)</u> . If an executive summary is not present, mark “N/A.”
<b>Program/Project Background</b>	
2. Are the basic characteristics of the project or program described (title, dates, funding organization, budget, implementing organization, location/map, target group)?	The project description plays a critical role in enabling the reader to understand the context of the evaluation, and involves several characteristics such as the title, dates, funding organization, budget, implementing organization, location/map, and target group. All of these characteristics play an important role and virtually all should be present to receive credit for this item in order to take a holistic view of whether the project is sufficiently well described. If one or two characteristics are missing or weak but you get the gist of the project and can answer all future questions, then check “yes.”
3. Is the project or program’s “theory of change” described (intended results (in particular the project Purpose); development hypotheses; assumptions)	The “theory of change” describes, via narrative and/or graphic depiction of the intended results and causal logic, how anticipated results will be achieved. You may see this described as the development hypotheses and assumptions underlying the project or program. We expect that a clear explanation of the theory of change/development hypotheses will be presented in the evaluation report <u>before</u> the evaluation’s findings are presented.
<b>Evaluation Purpose</b>	
4. Does the evaluation purpose identify the management reason(s) for undertaking the evaluation?	Evaluation policy states that USAID is conducting evaluations for learning and accountability purposes. Beyond that, it is important that the evaluation purpose identifies the specific decisions or actions the evaluation is expected to inform (e.g., continue, terminate, expand, or redesign an intervention). If a statement of the evaluation purpose is not found, or is only present in the SOW, mark “N/A.”
<b>Evaluation Questions</b>	
5. Are the evaluation questions clearly related to the evaluation purpose?	The evaluation questions, as stated in the evaluation report, should have a direct and clear relationship to the stated evaluation purpose. If no evaluation questions are provided in the body of the report before the findings, or in the SOW, check “N/A.” Even if questions are provided, this question cannot be answered if no evaluation purpose was included. Thus if item (4) above indicated that there was no purpose stated, then this question must be marked “N/A.”
6. Are the evaluation questions in the report identical to the evaluation questions in the SOW?	This question is about evaluation questions found in the body of the report <u>and</u> in the SOW. There must be questions in both places in order to address this question. If questions are present in only one of these two places, mark “N/A.”

<p>7. If the questions in the body of the report and those found in the SOW differ, does the report (or annexes) state that there was written approval for changes in the evaluation questions?</p>	<p>The evaluation SOW is the contract evaluators work from, so it is imperative that the questions/issues in the body of the evaluation report match those included in the SOW word for word. If the evaluation team changed, removed, or added evaluation questions/issues, USAID policy states that they should only have done so with written approval from USAID. While this written approval does not need to be included in an annex, it does need to be mentioned in the body of the report. If the answer to 6 is “yes” or “N/A” then mark 7 as “N/A.” If the answer to 6 is “no” then answer 7 with a “yes” or “no.”</p>
<p><b>Methodology</b></p>	
<p>8. Does the report (or methods annex) describe <u>specific data</u> collection methods the team used?</p>	<p>USAID requires that an evaluation report identify the data collection methods used, but does not indicate where this information must be presented. It is common to include the methodology description in the body of the report with a longer and more detailed methods annex, so be sure and check the annex. To receive credit, the methods description must be specific on how and from whom data will be collected. It is insufficient to say, “interviews will be conducted.” To be adequate a description of methods must indicate what types of interviews, estimated numbers, and with whom they will be conducted (e.g., key informant interviews, individual interviews with beneficiaries, group interviews).</p>
<p>9. Are the data collection methods presented (in the report or methods annex) in a manner that makes it clear which specific methods are used to address <u>each</u> evaluation question (e.g., matrix of questions by methods)?</p>	<p>USAID How-To guidance on evaluations advises that data collection methods should be explained in relation to each evaluation question/issue the evaluation team addressed. This information may be found within the body of the report or may be presented in a methods or design annex. While the methods can be associated to questions in a variety of ways, some evaluations use a matrix for this purpose that lists an evaluation question and then describes the data sources, data collection methods, sampling strategies, and data analysis methods. If no data collection methods are provided, or if no questions/issues exist, check the box for “N/A.”</p>
<p>10. Does the report (or methods annex) describe <u>specific data</u> analysis methods the team used? (frequency distributions; cross tabulations; correlation; reanalysis of secondary data)</p>	<p>USAID requires that an evaluation report identify the data analysis methods used, but does not indicate where this information must be presented. It is common to include the methodology description in the body of the report with a longer and more detailed methods annex. To receive credit, the data analysis methods description must be <u>specific about how, or through what method, data will be analyzed</u>. It is insufficient to say, “qualitative and quantitative analyses will be conducted” and instead must provide detailed information on the kinds of analyses to be conducted (e.g., frequency distributions, cross-tabs, correlations, content analysis, pattern analysis).</p>
<p>11. Are the data analysis methods presented (in the report or methods annex) in a manner that makes it clear how they are associated with the evaluation questions or specific data collection methods?</p>	<p>The evaluation report should make it clear which data analysis methods described were used to analyze data to answer specific evaluation questions/issues. [The question parallels #9 above for data collection methods.] Information on data analysis methods may be available within the body of the report or may be found in a methods or design annex. As indicated under item (9), some reports include a matrix that describes data analysis approaches as well as data collection methods in relation to each evaluation question. Note that wherever a discussion of data analysis methods takes place, it is acceptable for this description to relate data analysis methods to data collection methods, instead of directly to evaluation questions. If no data analysis methods are provided (marked “no” for previous question, #9), or if no questions exist, check the box for “N/A.”</p>
<p><b>Team Composition</b></p>	

<p>12. Did the report (or methods annex) indicate that the evaluation team leader was external to USAID?</p>	<p>USAID counts an evaluation as being external if the team leader is external, meaning that the team leader is an independent expert from outside of USAID <b>who has no fiduciary relationship with the implementing partner</b>. If the evaluation is a self-evaluation (USAID or its Implementing Partner is evaluating their own project/activity) then this answer must be no. To receive credit, the evaluation must indicate the team leader in either the body of the report (including cover or title page) or in the methods section. A search for the term “team leader” may expedite this process. <u>If the report is not explicit in stating the team leader was external, it may be inferred from a description of the team leader or the organization with which they are associated (e.g., university professor or evaluation firm that is not the project implementer)</u>. Independence may also be confirmed via a “no-conflict of interest” statement often included as an annex. If the report identifies that the team was independent, but there is no designated team leader, check “N/A.”</p>
<p>13. Did the report (or methods annex) identify at least one evaluation specialist on the team?</p>	<p>At least one member of the evaluation team must be an evaluation specialist and clearly indicated as such in either the body of the report or in the methods annex. The term “evaluation specialist” must be explicit and not implied.</p>
<p>14. Did the report (or methods annex) identify local evaluation team members?</p>	<p>USAID encourages the participation of country nationals on evaluation teams. The report need not use the word “local” specifically, but can be referred to by designation such as “Brazilian education specialist,” if in Brazil. This person could be any country national, including a foreign service national (FSN). Simply guessing a person’s country of origin based on their name is insufficient. Do not guess.</p>
<p>15. Did the report indicate that team members had signed Conflict of Interest forms or letters (check if the report says this or the COI forms are included in an annex)?</p>	<p>USAID requires that evaluation team members certify their independence by signing statements indicating that they have no conflict of interest or fiduciary involvement with the project or program they will evaluate. USAID guidance includes a sample Conflict of Interest form. It is expected that an evaluation will indicate that such forms, or their equivalent, are on file and available or are provided in an evaluation annex.</p>
<p><b>Study Limitations</b></p>	
<p>16. Does the report include a description of study limitations (lack of baseline data; selection bias as to sites, interviewees, comparison groups; seasonal unavailability of key informants)?</p>	<p>It is common for evaluators to encounter unexpected interferences with anticipated study designs such as unavailability of key informants or lack of access to activity sites. In other instances, stakeholder preferences may introduce selection biases. In any such instance, evaluators are obligated to include these “study limitations” and a description of the impact they have had on the evaluation. Study limitations may only be included for this item if they directly impact the evaluator’s ability to credibly and effectively answer an evaluation question (i.e., if all data can still be collected, even if inconveniently or at a higher cost, it is not a limitation). Limitations do not need to have their own distinct section provided they are located towards the end of the methodology description and before the introduction of findings.</p>
<p><b>Report Structure Responsiveness to Evaluation Questions</b></p>	
<p>17. Is the evaluation report structured to present findings in relation to evaluation questions, as opposed to presenting information in</p>	<p>The most straightforward way to meet USAID’s requirement that every evaluation question/issue be addressed, is a question-by-question (or issue-by issue) report structure. Historically, evaluations have not always taken this approach, and instead structured the report around such things as project objectives, or locations. If no evaluation questions/issues exist around which a report could be</p>

relation to project objectives or in some other format?	structured, check “N/A.” If the evaluation questions/issues and the team’s answers to those questions/issues are the dominant structure of the report, check “yes.”
18. Are <u>all</u> of the evaluation questions, including sub questions, answered primarily in the body of the report (as opposed to in an annex)	The purpose of an evaluation report is to provide the evaluators’ findings and recommendations on <u>each and every</u> evaluation question. Accordingly, USAID expects that the answers to <u>all</u> evaluation questions/issues, including any sub questions/issues, will be provided primarily in the body of the report. Answering main questions/issues in the body and sub-questions/issues in an annex is not consistent with USAID expectations. If no evaluation questions/issues are provided (either in the body of the report or in an annex) to which a team could respond, check “N/A.”
19. If any questions were not answered, did the report provide a reason why?	If the answer to question 18 is “yes,” mark this answer as “N/A.” If the answer to question 18 is “no,” does the evaluation report provide an explanation as to why specific questions were not answered or were answered somewhere other than in the body of the report?
<b>Findings</b>	
20. Did the findings presented appear to be drawn from social science data collection and analysis methods the team described in study methodology (including secondary data assembled or reanalyzed)?	USAID’s commitment to evidence-based decision-making is necessitating a shift to stronger and more replicable approaches to gathering data and presenting action recommendations to the agency. The more consistent use of credible social science data collection and analysis methods in evaluations is an important step in that direction (e.g., structured and well documented interviews, observation protocols, survey research methods). If the report did not describe the data collection and analysis methods used, check “N/A.”
21. For the findings presented within the evaluation report is there a transparent connection to the source(s) of the data? (60% of the beneficiaries interviews reported that...; reanalysis of school records shows...; responses from mayors indicate that...)	While most evaluation reports present sets of findings, it is not always clear where those findings came from. It is helpful to the reader to connect the sources of data to the findings those data are being used to support. For example, “children’s consumption of protein increased” does not indicate where that finding came from. Alternatively, “60% of mothers who participated in the survey stated that their children’s consumption of protein had increased” does a good job of connecting the finding to the source. This is true for both qualitative and quantitative findings. If the findings in the report were connected to sources of data as indicated above, check “yes.” If findings are generally presented without reference to their source, check “no.”
22. In the presentation of findings, did the team draw on data from the range of methods they used rather than answer using data from or primarily one method?	In addressing this question, only include those methods specifically referenced in the methods section of the report or in the methods annex. Of the methods actually used, the evaluation should demonstrate a balanced use of data from all data collection methods. If no methodologies were introduced from which they could later be drawn on, check “N/A.”
23. Are findings clearly distinguished from conclusions and recommendations in the report, at least by the use of language that signals transitions (“the evaluation found that...” or “the team concluded that...”)?	As defined by the evaluation policy, evaluation findings are “based on facts, evidence, and data...[and] should be specific, concise, and supported by quantitative and qualitative information that is reliable, valid, and generalizable”. The presence of opinions, conclusions, and/or recommendations mixed in with the descriptions of findings reduces a finding’s ability to meet USAID’s definition.

<p>24. Are quantitative findings reported precisely, i.e., as specific numbers or percentages rather than general statements like “some,” “many,” or “most”?</p>	<p>When presenting quantitative findings it is important to be precise so that the reader knows exactly how to interpret the findings and is able to determine the accuracy of the conclusions drawn by the evaluators. Precision implies the use of specific numbers and/or percentages as opposed to general statements like “some,” “many,” or “most.” If no potentially quantitative findings are provided, check “N/A.”</p>
<p>25. Does the report present findings about unplanned/ unanticipated results?</p>	<p>While evaluators may be asked to look for unplanned or unanticipated results in an evaluation question, it is common to come across such results unexpectedly. If such results are found, by request or unexpectedly, they should be included in the report.</p>
<p>26. Does the report discuss alternative possible causes of results/ outcomes it documents?</p>	<p>Though evaluators may be asked to look for alternative causes of documented results or outcomes in an evaluation question, it is possible for evaluators to come across such potential alternative causes unexpectedly. If any such causes are found, it is important that the evaluators bring such information to the attention of USAID.</p>
<p>27. Are evaluation findings disaggregated by sex at ALL levels (activity, outputs, outcomes) when data are person-focused?</p>	<p>The evaluation policy and USAID in general are making a big push for gathering sex-disaggregated data whenever possible. To support this focus, it is valuable for evaluators to include data collection and analysis methods that enable sex-disaggregation whenever the data they anticipate working with will be person-focused. Such data should be represented at all project levels from activities to outputs to outcomes to the extent possible. If no person-focused data was collected and therefore there was no data that could be disaggregated by sex, check “N/A.”</p>
<p>28. Does the report explain whether access/participation and/or outcomes/benefits were different for men and women when data are person-focused?</p>	<p>USAID expects that evaluations will identify/discuss/explain how men and women have participated in, and/or benefited from, the programs and projects it evaluates. This involves more than simply collecting data on a sex disaggregated basis. Addressing this issue can be presented in one general section or on a question-by-question basis; either is acceptable. If data was not collected in a person-focused manner for the evaluation, check “N/A.”</p>
<p><b>Recommendations</b></p>	
<p>29. Is the report’s presentation of recommendations limited to recommendations (free from repetition of information already presented or new findings not previously revealed)?</p>	<p>Presentation of recommendations in an evaluation report affects the usability of the report. Recommendations build on information previously introduced through findings and conclusions. Therefore, the presentation of recommendations does not need supporting findings and conclusions repeated or any new supporting findings or conclusions introduced. The presence of any information other than the specific, practical, and action oriented recommendations could have a diminishing effect on report usability. If no recommendations are present in the report, check “N/A.”</p>
<p>30. Do evaluation recommendations meet USAID policy expectations with respect to being specific (states what exactly is to be done, and possibly how)?</p>	<p>Recommendations that are specific are inherently more actionable than those which are not. The recommendation, “improve management of the project,” is much less specific than one that says “streamline the process for identifying and responding to clinic needs for supplies in order to reduce gaps in service delivery.” If no recommendations are presented in the evaluation report, check “N/A.”</p>
<p>31. Do evaluation recommendations meet USAID policy expectations with respect</p>	<p>USAID encourages evaluation teams to identify the parties who need to take action on each recommendation. Doing so makes it easier for USAID staff to understand and act on an evaluation’s implications. If no recommendations are presented in the evaluation report, check “N/A.”</p>

to being directed to a specific party?	
32. Are all the recommendations supported by the findings and conclusions presented (Can a reader can follow a transparent path from findings to conclusions to recommendations)?	Managers are more likely to adopt evaluation recommendations when those evaluations are based on credible empirical evidence and an analysis that transparently demonstrates why a specific recommendation is the soundest course of action. To this end, USAID encourages evaluators to present a clear progression from Findings <input type="checkbox"/> Conclusions <input type="checkbox"/> Recommendations in their reports, such that none of a report’s recommendations appear to lack grounding, or appear out of “thin air.” If no recommendations are presented in the evaluation report, check “N/A.”
<b>Annexes</b>	
33. Is the evaluation SOW included as an annex to the evaluation report?	This question checks on evaluation team responsiveness to USAID’s Evaluation Policy, Appendix I, requirement for including an evaluation SOW as an evaluation report annex.
34. Are sources of information that the evaluators used listed in annexes?	USAID’s Evaluation Policy, Appendix I, requires sources of information to be included as an evaluation report annex. Sources include both documents reviewed and individuals who have been interviewed. Generally it is not expected that names of survey respondents or focus group participants will be individually provided, as these individuals are generally exempted based on common/shared expectations about maintaining confidentiality with respect to individual respondents.
35. Are data collection instruments provided as evaluation report annexes?	This question focuses on the inclusion of data collection instruments in an evaluation annex including interview guides or survey questionnaires.
36. Is there a matching instrument for each and every data collection method the team reported that they used?	This question examines how comprehensive a set of the instruments used for collecting data for a USAID evaluation a report provides. USAID’s standard in its evaluation policy is “all” tools.
37. Were any “Statements of Differences” included as evaluation annexes (prepared by team members, or the Mission, or Implementing Partner, or other stakeholders)	Including “Statements of Differences” has long been a USAID evaluation report option. This question determines how frequently “Statements of Differences” are actually included in USAID evaluations. Statements are often written by evaluation team members, or alternatively by the Mission, a stakeholder, or implementing partner.
<b>Evaluation Data Warehousing</b>	
38. Does the evaluation report explain how the evaluation data will be transferred to USAID (survey data, focus group transcripts)?	USAID evaluation policy (p. 10) calls for the transfer of data sets from evaluations to USAID, so that, when appropriate, they can be reused in other assessment and evaluations. Given this requirement, it is helpful if an evaluation report indicates how and when that transfer was made.
<b>SOW Leading Indicator of Evaluation Quality (answer if SOW is a report annex)</b>	

<p>39. Does the evaluation SOW include a copy or the equivalent of Appendix I of the evaluation policy?</p>	<p>The 2011 USAID Evaluation Policy required that statements of work (SOWs) for evaluations include the language of Appendix I from the policy. This was followed by ADS 201maa “Criteria to Ensure the Quality of the Evaluation Report,” which states that the criteria must be included in every evaluation SOW. If no SOW is included as an annex to the evaluation report, check “N/A.”</p>
<p><b>Additional Questions About Basic Evaluation Characteristics</b></p>	
<p>40. Does the report include a table of contents?</p>	<p>Including a table of contents informs the reader on what the report covers and provides the reader with page numbers to better access information in a given section. Ideally a table of tables and/or a table of figures will also be included to facilitate access to data.</p>
<p>41. Does the report include a glossary and/or list of acronyms?</p>	<p>A high-quality evaluation report should include a glossary and/or a list of acronyms used throughout the report since not all readers are familiar with the acronyms, abbreviations, or nuanced language specific to a given subject or country.</p>
<p>42. Is the report well-written (clear sentences, reasonable length paragraphs) and mostly free of typos and other grammatical errors?</p>	<p>High-quality evaluation reports give the appearance of having been edited or peer-reviewed to remove any grammatical, syntax, or punctuation inconsistencies or errors. Attempting to read an evaluation report that contains errors, inconsistencies, or unclear sentences prevents the reader from being able to digest or comprehend the content of the report.</p>
<p>43. Is the report well-organized (each topic is clearly delineated, subheadings used for easy reading)?</p>	<p>A high-quality evaluation report should be well-organized to facilitate ease of reading and ability for the reader to digest the content of the report in a logical manner. The use of section headings, sub-headings, and titles breaks up what may be long and dense sections of reports.</p>
<p>44. Is the date of the report given?</p>	<p>The date of the report should be included in the report or on the front cover of the report. This may be the date submitted to or approved by USAID, or the date disseminated to the public.</p>
<p>45. Is the name of the evaluation team leader present in the report or on the report cover?</p>	<p>The names and roles of all team members should be included either in the body of the report or on the front cover. At very least the evaluation team leader must be readily identified by name as they are the person responsible for the final report deliverable.</p>

## Evaluation Descriptive Data and Content Tool

#	Question	Data Type	Response
1	a) What is the <u>scope</u> of the evaluation? <ul style="list-style-type: none"> <li>• Single activity (one country)</li> <li>• Project-level (one country)</li> <li>• DO level (one country)</li> <li>• Sector-wide (one country)</li> <li>• Other multiple activities (one country)</li> <li>• Single activity (multiple countries)</li> <li>• Multiple activities (multiple countries)</li> <li>• Regional program or project (multiple countries)</li> <li>• Global program or project (funded by USAID/W)</li> <li>• Other</li> </ul>	Select one	
	b) Is this an <u>integrated</u> activity or project (also called <u>multisectoral</u> ) covering two or more sectors? e.g. project or activity jointly funded by two or more USAID Technical Offices such as Education and Health	Y / N / unclear, flag for further review	
2	a) Are the <u>questions</u> numbered?	Y / N / N/A	
	b) Highest number assigned, not counting sub-questions	#	
	c) Count of all question marks, including sub-questions	#	
	d) Count of all questions, including compound questions	#	
3	a) <b>For performance evaluations only:</b> Did the evaluation include any cause-and-effect question that asks about <u>causality/attribution</u> ? (select N/A if reviewing an impact evaluation)	Y / N / N/A	
	b) If so, which type of design did the evaluation use to examine causality/attribution? <ul style="list-style-type: none"> <li>• Experimental design</li> <li>• Quasi-experimental design</li> <li>• Non-experimental approach</li> </ul>	Select one	

#	Question	Data Type	Response		
	<ul style="list-style-type: none"> <li>None: While there were questions about causality/attribution in the list, no overall design for answering these questions was presented.</li> <li>N/A - no questions about causality/attribution</li> <li>N/A - impact evaluation</li> </ul>				
	c) If a design is identified, describe here	text			
4	a) Which <u>data collection methods</u> did the evaluation report say it would use in the methodology and which did it demonstrate use of in the report?	Check all that are "yes"	Planned use	Visible use	
	<ul style="list-style-type: none"> <li>Cull data from document review / secondary source data</li> </ul>				
	<ul style="list-style-type: none"> <li>Cull facts from activity performance monitoring data</li> </ul>				
	<ul style="list-style-type: none"> <li>Structured observation</li> </ul>				
	<ul style="list-style-type: none"> <li>Unstructured observation</li> </ul>				
	<ul style="list-style-type: none"> <li>Individual interviews</li> </ul>				
	<ul style="list-style-type: none"> <li>Group interviews or focus groups</li> </ul>				
	<ul style="list-style-type: none"> <li>Survey</li> </ul>				
	b) For all types of data collection planned, indicate if they were conducted in person or virtual	Check all that are "yes"	In person	Virtual	Unclear
	<ul style="list-style-type: none"> <li>Individual interviews</li> </ul>				
<ul style="list-style-type: none"> <li>Group interviews or focus group</li> </ul>					
<ul style="list-style-type: none"> <li>Survey</li> </ul>					
5	Which <u>data analysis methods</u> did the evaluation report say it would use in the methodology and which did it demonstrate use of in the report?	Check all that are "yes"	Planned use	Visible use	
	<ul style="list-style-type: none"> <li>Descriptive statistics (frequency, percent, ratio, cross-tabulations)</li> </ul>				
	<ul style="list-style-type: none"> <li>Inferential statistics (regression, correlation, t-test, chi-square)</li> </ul>				
	<ul style="list-style-type: none"> <li>Content or pattern analysis (describes patterns in qualitative responses)</li> </ul>				

#	Question	Data Type	Response
6	Did the evaluation report discuss any <u>gender-sensitive data approaches</u> to data collection, either in designing the instruments or protocols?	Y / N	
7	a) Did the activity include <u>youth skill building</u> and/or other capacity strengthening support with the <u>aim of increasing youth employment</u> ?	Y / N	
	b) If yes to 7.a, does the report state that the activity's approach to youth skill building and/or other capacity strengthening was associated with an increase in employment?	Y / N / N/A	
	c) If yes to 7.b, provide text	text	
	d) Did the project/activity address <u>youth civic engagement and/or leadership</u> ?	Y / N	
8	Did the evaluation report state that <u>beneficiaries participated</u> in any of these stages of the evaluation?	Check all that are "yes"	-
	• Evaluation design/methods selection		
	• Data collection		
	• Data analysis		
	• Report writing		
	• Formulation of recommendations		
	• Other type of participation		
	• Evaluation stated beneficiaries participated, but did not describe when		
9	Did the evaluation report include any <u>visual representations of data</u> in the body of the report?	Y / N	
10	If the evaluation report included a section on <u>lessons learned</u> , provide text.	text	
11	If the evaluation report described any aspects of the activity as " <u>innovative</u> ", provide text.	text	
12	Successes and challenges - For each topic below, provide any text from the evaluation report related to success or failure in project design, implementation, and/or performance.	text	-
	a) Theory of Change and assumptions		

#	Question	Data Type	Response
	b) Monitoring system, learning, and adaptive management		
	c) Performance target setting, goals, and objectives		
	d) Beneficiary selection and targeting ( <i>individuals and organizations/institutions</i> )		
	e) Government engagement ( <i>at any level in country</i> )		
13	If “YES” to Q25 on the quality questionnaire (“Does the report present findings about unplanned/unanticipated results?”), provide text	text	
14	a) Did the evaluation report address the sustainability of the activity’s <u>outcomes</u> ? If yes, provide text.	text	
	b) Did the evaluation report address the sustainability of the activity’s <u>interventions</u> ? If yes, provide text.	text	
	c) According to the evaluation report, was a <u>sustainability plan</u> developed during the <u>activity design</u> or at the beginning of the activity? <ul style="list-style-type: none"> <li>• Yes, the evaluation stated that the plan was developed</li> <li>• No, the evaluation stated that the plan was not developed</li> <li>• Plan not addressed in the evaluation</li> </ul>	Select one	
	d) If yes or no to 14.c, provide text	text	
	e) According to the evaluation report, was a <u>sustainability plan</u> developed at the <u>end</u> of the activity? <ul style="list-style-type: none"> <li>• Yes, the evaluation stated that the plan was developed</li> <li>• No, the evaluation stated that the plan was not developed</li> <li>• Plan not addressed in the evaluation</li> </ul>	Select one	
	f) If yes or no to 14.e., provide text	text	
15	a) If the evaluation addressed any effects of <u>COVID-19</u> on the <u>activity</u> , provide text	text	
	b) If the evaluation addressed any effects of <u>COVID-19</u> on the <u>evaluation</u> , provide text	text	

#	Question	Data Type	Response
16	Provide the text of all recommendations. If the recommendations are more than two pages of text or in a format that is difficult to copy/paste (i.e. a table), provide the page number.	text	
17	Provide the text of the <u>evaluation questions</u> .	text	
18	a. If you as the reader felt that this <u>activity</u> made a significant contribution to a development outcome, please provide a note here.	text	
	b. If you as the reader felt that this <u>evaluation</u> was well done in any way that makes it a good example to highlight for future learning, please provide a note here.	text	
19	Any other comments or flags for this evaluation	text	

## Evaluation Descriptive Data and Content Tool - Rater's Guide

#	Question	Guide
1	<p>a) What is the <u>scope</u> of the evaluation?</p> <ul style="list-style-type: none"> <li>• Single activity (one country)</li> <li>• Project-level (one country)- explicitly examines two or more activities under a project. Also called whole-of-project evaluation-</li> <li>• DO level (one country) – explicitly examines all elements under a USAID Development Objective (DO), e.g., “economic growth improved”, “food security increased”</li> <li>• Sector-wide (one country) – e.g., all agriculture, all health projects/activities</li> <li>• Integrated activity or project (also called multisectoral) covering two or more sectors- Project or activity jointly funded by two or more USAID Technical Offices (e.g. Education and Health; Education, Health, and Economic Growth)</li> <li>• Other multiple activities (one country) evaluation, e.g., several activities in one district, or several activities focused on youth employment</li> <li>• Single activity (multiple countries) e.g., approach to sexual violence in schools in Ghana and Malawi</li> <li>• Multiple activities (multiple countries), e.g., worldwide review of Mission funded trade projects</li> <li>• Regional program or project (funded by a Regional Mission/Office or Bureau); e.g., Mekong River cooperation project involving multiple countries</li> <li>• Global program or project (funded by USAID/W), e.g., worldwide assistance to missions on gender assessments</li> <li>• Other</li> </ul>	<p>This item refers to what exactly was being evaluated. Evaluations can look at individual projects or can look at multiple projects at a time and they can focus on an individual country or a group of countries. It is important for our purposes to be able to distinguish evaluations based on their scope. Some of the scopes provided are fairly straightforward while others are a bit more nuanced and are given more detail below.</p> <p>An evaluation of a <b>single project or activity</b> corresponds to one implementing mechanism (contract, grant, cooperative agreement), regardless of the number of subcontractors or tasks/activities within that implementing mechanism.</p> <p>When evaluating multiple projects <b>within one country</b> there are three options:</p> <ul style="list-style-type: none"> <li>• A program-level evaluation would explicitly examine every element within one of the country mission’s Development Objectives (DOs). DOs focus on large technical issues such as economic growth or food security and would encompass all elements that contribute to achieving the DO.</li> <li>• A sector-wide evaluation would look at all, or a sample of, the projects within a given technical sector such as agriculture or education. This may crosscut or be a subset of a DO.</li> <li>• The category “other multi-project single-country” might focus on all, or a sample of, the projects within a geographic region of a country or a group of activities, for example, focused on youth employment.</li> </ul> <p>When evaluating <b>across multiple countries</b>, there are four options:</p> <ul style="list-style-type: none"> <li>• An example of a single-project multi-country evaluation might focus on an approach to dealing with sexual violence in schools in Malawi and Ghana</li> </ul>

#	Question	Guide
		<ul style="list-style-type: none"> <li>• An example of a multi-project multi-country evaluation might focus on a sample of Mission-funded trade projects around the world</li> <li>• A regional program or project evaluation is one that is funded by a regional office or bureau and is focused on a specific geographic region or group of countries. For example, climate change along the Mekong River.</li> <li>• A global project is funded through USAID/Washington. For example, a project that can help any mission do a gender assessment.</li> </ul>
	b) Is this an <u>integrated</u> activity or project (also called <u>multisectoral</u> ) covering two or more sectors? e.g. project or activity jointly funded by two or more USAID Technical Offices such as Education and Health	An integrated activity or project is funded by two or more sectors. Any type of activity or project can be integrated. If unclear, flag for further review by the Team Leader.
2	a) Are the <u>questions</u> numbered?	This is a yes/no question about whether questions (not issues) found in the body of the report, or in the SOW if there were none in the body of the report, had been assigned numbers. If there are questions in both the body of the report and the SOW, the <b>questions in the body of the report</b> take precedence in terms of answering all elements of this set of questions.
	b) Highest number assigned, not counting sub-questions	In the simplest instance, questions would be numbered 1-5. If there are sub-questions, (e.g., 5a, 5b) then the highest number of questions would still be 5. In other instances, questions might be in groups (e.g., A, 1-5, and then B, 1-6). In this type of case the number of numbered questions would be 11. If you answered “no” on 17 (a) above, enter 0 (zero) for 17 (b)
	c) Count of all question marks, including sub-questions	This is a simple count of how many question marks were used in presenting the questions in the body of the report, or in the SOW if no questions were found in the body of the report. Don’t worry about hidden or compound questions, just count question marks. If there are questions with no question marks, they cannot be counted, only questions with question marks.

#	Question	Guide
	d) Count of all questions, including compound questions	For this item, we are looking for a count of all questions beyond those distinguished by a question mark. Compound, or hidden questions, are questions with an “and” in them or perhaps a list of items an evaluator is being asked to look at within a specific question. An example of this might be, “Did crop yields increase and was community resilience strengthened?”
3	a) <b>For performance evaluations only:</b> Did the evaluation include any cause-and-effect questions that ask about <u>causality/attribution</u> ?	If the evaluation team is responsible for answering one or more questions or issues that ask about causality or attribution, pick “yes”. If they were not, pick “no”. If they were not asked to answer any questions or if this is an impact evaluation, pick “N/A”.
	b) If so, which type of design did the evaluation use to examine causality/attribution? <ul style="list-style-type: none"> <li>• Experimental design</li> <li>• Quasi-experimental design</li> <li>• Non-experimental approach</li> <li>• None: While there were questions about causality/attribution in the list, no overall design for answering these questions was presented.</li> <li>• N/A - no questions about causality/attribution</li> <li>• N/A - impact evaluation</li> </ul>	For questions or issues of causality and attribution, there are three categories of evaluation designs to choose from: <ul style="list-style-type: none"> <li>• <b>Experimental design:</b> The evaluation report says it used an experimental design or provided equivalent words (control group, randomized assignment, randomized control trial)</li> <li>• <b>Quasi-experimental design:</b> The evaluation report says it used a quasi-experimental design or provided equivalent words (comparison group, regression discontinuity, matching design, propensity score matching, interrupted time series)</li> <li>• <b>Non-experimental approach:</b> The evaluation report says it used a specific non-experimental approach for examining causality or attribution (outcome mapping; identification &amp; elimination of alternative possible causes (modus operandi); contribution analysis, case study).</li> </ul> In order to fall into one of these categories the evaluation design must be specifically discussed in the body of the evaluation report. If not discussed, or if discussed exclusively in an annex exclusively, please select “none.” The two N/A options should only be used when the response to 3.a) was “no” or “N/A”.
4		Data Collection Methods (team said it planned to use) – For this item, we are looking for every data collection method that the evaluation team

#	Question	Guide
	<p>a) Which <u>data collection methods</u> did the evaluation report say it would use in the methodology and which did it demonstrate use of in the report?</p> <ul style="list-style-type: none"> <li>• Cull data from document review / secondary source data</li> <li>• Cull facts from activity performance monitoring data</li> <li>• Structured observation</li> <li>• Unstructured observation</li> <li>• Individual interviews</li> <li>• Group interviews or focus groups</li> <li>• Survey</li> </ul>	<p>stated that they planned to use (either in the body of the report or in a methodology annex). In the instance that the data collection team introduces a data collection method, but misstates what the method actually is, and there is enough information provided for you as a coder to appropriately re-categorize it, please do so (e.g., if an evaluation claims to be doing qualitative interviews, but the description and a look at the data collection instrument indicate that it is actually a survey, mark it as a survey). An evaluation can use more than one data collection method. A list of data collection methods based on the most common methods used in previous studies are shown on the demographic sheet. Select all options that apply.</p> <p>Data Collection Methods (data actually used) - For this item, we are looking for the presentation of data that shows which data collection methods were actually used. For example, “20% of the survey respondents said” indicates that the survey method was actually used. The demographic sheet shows the same list of data collection methods as you saw in item 19. For every method you mark that they planned to use, look to see if there was data linked to words about the method that would indicate it was actually used. Select all options that apply.</p>
	<p>b) For all types of data collection planned, indicate if they were conducted in person or virtual</p> <ul style="list-style-type: none"> <li>• Individual interviews</li> <li>• Group interviews or focus groups</li> <li>• Survey</li> </ul>	<p>For this item, we are looking for whether data was collected in person or virtually for every data collection method that the evaluation team stated that they planned to use. Do not guess - if the data collection mode is not explicitly stated, select “unclear”. Check all that apply, for example an evaluation could include both an in person and virtual survey.</p>
5	<p>Which <u>data analysis methods</u> did the evaluation report say it would use in the methodology and which did it demonstrate use of in the report?</p>	<p>Data Analysis Methods (team said it planned to use) – For this item, we are looking for every data analysis method that the evaluation team stated that they planned to use (either in the body of the report or in a methodology annex). An evaluation can use more than one data analysis method. A list of data analysis methods based on the most common</p>
	<p>Descriptive statistics (frequency, percent, ratio, cross-tabulations)</p>	

#	Question	Guide
	Inferential statistics (regression, correlation, t-test, chi-square)	methods used in previous studies are shown on the demographic sheet. Select all that apply.
	Content or pattern analysis (describes patterns in qualitative responses)	Data Analysis Methods (data actually used) - For this item, we are looking for the presentation of data that shows which data analysis methods were actually used. For every method you mark that they planned to use, look to see if there was analysis language, tables, or graphs that would indicate it was actually used. Select all that apply.
6	Did the evaluation report discuss any <u>gender-sensitive data approaches</u> to data collection, either in designing the instruments or protocols?	This item is to capture whether the report or the methods annex discussed any considerations for gender sensitivity in the design of the instruments (i.e. gender specific questions or additional lines of inquiry) or the data collection protocols (i.e. using male or female data collectors, gender composition of group interviews and/or focus groups, etc.)
7	a) Did the activity include <u>youth skill building</u> and/or other capacity strengthening support with the <u>aim of increasing youth employment</u> ?	This item captures examples of interventions that were successful in increasing youth employment through youth skill building and/or other youth capacity strengthening support.
	b) Does the report state that the activity's approach to youth skill building and/or other capacity strengthening was associated with an increase in employment?	For 7.c), provide a summary of the relevant text and all page numbers.
	c) If yes to 7.b, provide text	
	d) Did the project/activity address <u>youth civic engagement and/or leadership</u> ?	This item captures whether an intervention was tasked with addressing youth civic engagement and/or leadership.
8	Did the evaluation report state that <u>beneficiaries participated</u> in any of these stages of the evaluation? <ul style="list-style-type: none"> <li>• Evaluation design/methods selection</li> <li>• Data collection</li> <li>• Data analysis</li> <li>• Report writing</li> <li>• Formulation of recommendations</li> <li>• Other type of participation</li> </ul>	This item is to capture any phase during which beneficiaries participated as described in either the body of the report or methodology section. Select all that apply.

#	Question	Guide
	<ul style="list-style-type: none"> <li>Evaluation stated that beneficiaries participated, but did not describe when</li> </ul>	
9	Did the evaluation report include any visual representations of data (i.e. charts, infographics, etc.)	This item captures whether or not the evaluation team produced visual representations of data. This asks for presence/absence of graphics, not for any judgment call regarding their quality.
10	If the evaluation report included a section on <u>lessons learned</u> , provide text.	These should be identified by the evaluation as “lessons learned”, either in a distinct section of the report or in the conclusions. Do not make any value judgments as to whether they are actually lessons learned. Provide a summary of the relevant text and all page numbers.
11	If the evaluation report described any aspects of the activity as “ <u>innovative</u> ”, provide text.	These should be practices identified in the evaluation report as “innovation”, “innovative”, etc. Do not make any judgments as to whether it is actually an innovation. Provide a summary of the relevant text and all page numbers.
12	Successes and challenges - For each topic below, provide any text from the evaluation report related to success or failure in project design, implementation, and/or performance.	These series of questions aim to capture information related to successes and/or failures in project design, implementation, and/or performance across topics of interest. Provide a summary of the relevant text and all page numbers.
	a) Theory of change and assumptions	This topic could include activities where the theory of change and/or assumptions either proved to be correct or incorrect during implementation. Of particular interest is whether the Theory of Change articulated during project design was found to hold throughout the activity.
	b) Monitoring system, learning, and adaptive management	Examples could include: <ul style="list-style-type: none"> <li>indicator quality</li> <li>data quality</li> <li>data storage</li> <li>data analysis</li> <li>learning and adaptive management</li> </ul>
	c) Performance target setting, goals, and objective	This topic specifically addresses issues with target setting such as targets being set too high during project design or inception. Could also include

#	Question	Guide
		issues with goals and/or objectives being unrealistic - either too high or too low.
	d) Beneficiary selection and targeting	Successes and challenges with how the activity selected and targeted beneficiaries. Could include issues with methodology, access, etc. Beneficiaries can include individuals, groups of individuals, organizations, institutions, etc.
	e) Government engagement	Successes and challenges with working with the host country government at all levels, including local, regional, and national.
13	If “YES” to Q25 on the quality questionnaire (“Does the report present findings about unplanned/unanticipated results?”), provide text	This item captures examples of unplanned/unanticipated results. For example, an activity provided teacher training with the planned result of increasing literacy, but the evaluation also found that training was correlated with increased teacher retention. “Results” can be either positive or negative.
14	a) Did the evaluation report describe the sustainability of the activity’s <u>outcomes</u> ?	This item captures whether the evaluation explicitly discussed the sustainability of the activity’s outcomes, and can be either positive or negative. Provide a summary of the relevant text and all page numbers.
	b) Did the evaluation report describe the sustainability of the activity’s <u>interventions</u> ?	
	c) According to the evaluation report, was a <u>sustainability plan</u> developed during the <u>activity design</u> or at the beginning of the activity?	This item captures whether the evaluation explicitly discussed whether the activity had a sustainability plan at the start and end of the activity. This can also be described as an “exit strategy” if the strategy specifically addresses sustainability. <ul style="list-style-type: none"> <li>• Yes, the evaluation specifically stated that the plan <u>was</u> developed</li> <li>• No, the evaluation specifically stated that the plan <u>was not</u> developed</li> <li>• Plan not addressed in the evaluation - a sustainability plan was not discussed in the evaluation at all.</li> </ul> Provide a summary of the relevant text and all page numbers - both positives (yes, they had a plan) and negatives (no, they didn’t have a plan).
	d) If yes or no to 14.c, provide text	
	e) According to the evaluation report, was a <u>sustainability plan</u> developed at the <u>end</u> of the activity?	
	f) If yes or no to 14.e, provide text	

#	Question	Guide
15	a) If the evaluation addressed any effects of <u>COVID-19</u> on the <u>activity</u> , provide text	This item captures examples of how activities were affected by COVID-19, which could include changes or limitations in implementation, project design, beneficiary access, monitoring, etc. Provide a summary of the relevant text and all page numbers.
	b) If the evaluation addressed any effects of <u>COVID-19</u> on the <u>evaluation</u> , provide text	This item captures examples of how evaluations were affected by COVID-19, which could include changing or selecting specific data collection approaches, changes in timelines, changes in access, etc. Provide a summary of the relevant text and all page numbers.
16	Provide the text of all recommendations. If the recommendations are more than two pages of text or in a format that is difficult to copy/paste (i.e. a table), provide the page number.	This item captures the evaluation recommendations. Copy/paste the exact text of the recommendations, do not summarize.
17	Provide the text of the evaluation questions.	This item captures the evaluation questions. Copy/paste the exact text of the questions, do not summarize. Capture the questions from the body of the report (usually in the methodology section before the findings, conclusions, and recommendations). If there are no questions in the body of the report, copy them from the SOW. If no questions in either place, enter N/A.
18	a) If you as the reader felt that this activity made a significant contribution to a development outcome, please provide a note here.	The purpose of these two questions is to provide a set of examples for the Bureau to hold up as particularly high performing activities or well done evaluations. Identify the evaluations that you think meet this criteria.
	b) If you as the reader felt that this <u>evaluation</u> was well done in any way that makes it a good example to highlight for future learning, please provide a note here. (i.e. well structured, interesting data collection approach, high quality data visualization, etc.)	
19	Any other comments or flags for this evaluation	Open text field for any additional data capture or reviewer comments.