USAID
FROM THE AMERICAN PEOPLE

USAID
BUREAU FOR HUMANITARIAN ASSISTANCE
TECHNICAL GUIDANCE FOR MONITORING, EVALUATION, AND REPORTING FOR RESILIENCE FOOD SECURITY ACTIVITIES V2.0

Originally published May 2020
Updated May 2021

Record of Change: Technical Guidance for Monitoring, Evaluation, and Reporting for Resilience Food Security Activities

The following changes have been made to *Technical Guidance for Monitoring, Evaluation, and Reporting for Resilience Food Security Activities* since May 2020. Applicants must incorporate these changes into their applications. The most recent changes are listed first.

Date of Change	Section	Change
5/7/21	Entire document	Replacement of references to "FFP" and "Food for Peace" with "BHA" and "Bureau for Humanitarian Assistance"
5/7/21	Entire document	Replacement of references to "DFSA" and "development food security activities" with "RFSA" and "resilience food security activities"

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ADS	USAID Automated Directives System
AOR	Agreement Officer's Representative
ARR	Annual Results Report
ART	Award Results Tracking
AS	Annual or Routine Participant Survey
BHA	Bureau for Humanitarian Assistance
BL	Baseline
DDL	Development Data Library
DEC	USAID Development Experience Clearinghouse
DQA	Data Quality Assessment
EMMP	Environmental Mitigation and Monitoring Plan
FFS	Farmer Field School
FY	Fiscal Year
IEE	Initial Environmental Examination
IGA	Income-Generating Activity
IPTT	Indicator Performance Tracking Table
LOA	Life of The Award
LogFrame	Logical Framework
M&E	Monitoring and Evaluation
MCHN	Maternal and Child Health and Nutrition
MIS	Management Information System
MTE	Midterm Evaluation
PIRS	Performance Indicator Reference Sheet
PPS	Probability Proportional to Size
PREP	Pipeline and Resources Estimate Proposal
QuIPS	Qualitative Inquiry Planning Sheet
R	Required
RFSA	Resilience Food Security Activity
RiA	Required if Applicable
RCT	Randomized Control Trial
RF	Results Framework
RM	Routine Monitoring
R&I	Refine and Implement
SOW	Statement of Work
TOC	Theory of Change
USAID	U.S. Agency for International Development

BACKGROUND AND PURPOSE

The U.S. Agency for International Development's (USAID) Bureau for Humanitarian Assistance (BHA) funds resilience food security activities (RFSA) with an objective to improve and sustain the food and nutrition security of vulnerable populations. USAID applies the best available evidence to document activity effectiveness and improve program design and implementation. BHA has an obligation to the federal government and the American people to ensure that resources are used efficiently to achieve the best possible food security outcomes and that, in the process, food assistance actors learn from experience how to improve activity design and implementation.

This *BHA Technical Guidance for Monitoring, Evaluation, and Reporting for Resilience Food Security Activities* describes key monitoring, evaluation, and reporting responsibilities of BHA resilience activity applicants and partners. This guidance reflects BHA's approach to RFSA design and implementation called "Refine and Implement" (R&I), which enables partners and USAID to focus on formative research and/or analysis in the first phase of an activity life cycle, then further refine the technical approach(es) that will be carried out in the second phase of the activity life cycle. Also note that this document replaces the May 2020 M&E guidance for Development Food Security Activities produced by the legacy USAID/Office of Food for Peace.

Chapter 1 provides an overview of BHA M&E requirements and timelines for submission of the M&E Plan components. Chapter 2 outlines requirements for the Theory of Change (TOC), Logical Framework (LogFrame), Indicator Performance Tracking Table (IPTT), Performance Indicator Reference Sheets (PIRS), and Qualitative Inquiry Planning Sheets (QuIPS). Guidance on monitoring and M&E Staffing and Capacity Development Strategy is presented in Chapter 3, and guidance for evaluation plans is presented in Chapter 4.

Annex I provides an overview of the baseline study; this annex serves to inform partners about the objectives and methodology of the baseline. Annex II provides guidance on midterm evaluations; this annex is intended for activity implementing partners and research/evaluation partners who will be leading or participating in midterm evaluations. Annex III provides an overview of the interim/final evaluation; this annex serves to inform partners about the objectives and methodology of the interim/final evaluation. Annex IV provides technical guidance for partners on management information system (MIS) principles. Annex V is a template for Qualitative Inquiry Planning Sheets (QuIPS), which should be used by partners to outline performance monitoring and special studies using qualitative methods and tools. Applicants and partners are encouraged to use this and other templates included in this guidance document when developing their M&E Plans, but may use other formats if the required information is included.

CHAPTER 1. SUMMARY OF MONITORING AND EVALUATION PROCESSES

1.1 PURPOSE OF MONITORING AND EVALUATION

USAID relies on evidence collected through monitoring and evaluations to document and assess activity processes and outcomes, to learn, and to make decisions.

Monitoring is routine collection and analysis of quantitative or qualitative information throughout the life of the award (LOA) to verify that protocols are respected, implementation and outputs are on schedule, and evolving changes are consistent with the underlying theory of change (TOC).

Evaluations collect and analyze information about the characteristics, processes, and/or outcomes of interventions and are used for documenting performance and/or to inform decisions about current and/or future interventions. Evaluations provide opportunities to review both planned and unplanned results and to re-examine activity design. For BHA activities, midterm evaluations (MTEs) gather evidence on the effectiveness and efficiency of activity implementation and on the acceptability of activity implementation to participants, BHA, and the partner. Results are used to identify adjustments to implementation that would improve outcomes during the activity's remaining implementation period.

Final performance evaluations collect data on project outcomes to measure performance against baseline values, analyze successes and challenges of the activity, and inform future BHA and partner programming and learning. Note that BHA may choose to use impact evaluations with randomized control trial or other experimental or quasi-experimental designs in lieu of a standard final performance evaluation design. For BHA R&I activities, BHA will manage an interim evaluation in the fourth year of the award, which will serve to inform a possible cost extension of high-performing activities. If an activity is not extended, the interim evaluation will serve as the final performance evaluation; if an activity is extended, an additional final evaluation will be conducted in the final year of the extended award life (in addition to the interim evaluation in Year 4). The term "interim/final evaluation" is used in this document to refer to the performance or impact evaluation that will be conducted in the fourth year of the award.

1.2 MONITORING AND EVALUATION PLAN

A Monitoring and Evaluation (M&E) Plan is a road map for activity M&E implementation. Its primary purpose is to document an activity's M&E processes, methods, and elements in sufficient detail to enable all partner staff, especially new staff, to continue implementation of the M&E system after turnover of key M&E staff. An M&E Plan also demonstrates to BHA that an applicant or partner has a rigorous system for monitoring and evaluating activity performance that produces accurate, meaningful, and useful data for decision making.

The M&E Plan must include:

- A Theory of Change, which comprises a diagram and complementary narrative or table;
- A logical framework (LogFrame);
- An Indicator Performance Tracking Table (IPTT);
- Reference sheets, which includes Performance Indicator Reference Sheet (PIRS) and Qualitative Inquiry Planning Sheets (QuIPS);
- A Monitoring Strategy that includes a description of quantitative and qualitative data collection; a data flow diagram/matrix; a data quality assurance strategy; a data management and safeguarding plan; and a data utilization strategy;
- An M&E Staffing and Capacity Development Strategy; and
- An Evaluation Plan that describes the partner's participation in the baseline study; the general timing and scope of the midterm evaluation; and the partner's participation in the interim/final evaluation that will be conducted in the fourth year of implementation.

1.3 M&E REQUIREMENTS AND SUBMISSION TIMELINE

This section summarizes the key M&E expectation for applicants and current BHA resilience activity partners from application through the end of the activity. Table 1 summarizes the requirements at each stage of an activity. In addition, the activity award may include award-specific monitoring, evaluation, and reporting requirements. Partners should thoroughly review their award documents, and coordinate with the Agreement Officer's Representative (AOR) and relevant USAID Missions to ensure that they fulfill all requirements. Any questions about M&E requirements should be directed to the award AOR.

1.3.1 M&E REQUIREMENTS AT THE APPLICATION STAGE

BHA requires applicants to provide an abridged M&E Plan as part of every application. The abridged M&E Plan should include the following:

- A detailed TOC diagram and complementary narrative/table. The TOC should display the incremental, causal linkages from intervention to Goal. The TOC should reflect all interventions that are relevant to achieve the higher-level outcomes in the TOC regardless of which entities (i.e., the BHA implementing partner, another NGO, or local government) are implementing them. The TOC should also include assumptions and/or hypotheses about the operational, geographic, and cultural context.
- A LogFrame that includes all anticipated results that are within the manageable interest of the proposed activity. The LogFrame should include at least one indicator for each result and may include qualitative inquiries (i.e., qualitative monitoring or qualitative studies) if/where relevant. All BHA required (R) and required if applicable (RiA) indicators should be included and targets must be provided for all baseline (BL) indicators[1] at the Purpose- and Sub-Purpose-level. (Note that because the baseline values are not available at the application stage, the endline targets may be expressed in relation to the baseline value, e.g., "baseline plus 10 percentage points").

> **Note.** For indicators that are expressed as a prevalence or percent, such as % of children under 5 who are stunted, FFP requires targets to be reported as a change in percentage points, not percent change. If the applicant is aiming for a 2 percentage point reduction in stunting annually over 5 years, the LogFrame should state "Baseline value minus 10 percentage points" at application.

[1] Note that BHA uses the acronym "BL" to refer to baseline indicators collected during the baseline study by the research/evaluation partner and are collected at the population level for pre-post performance evaluation designs and may be either at the population- or participant-level for impact evaluation designs. These indicators will also be collected in the quantitative survey conducted as part of the interim/final evaluation.

- A basic overview of the planned monitoring strategy, including a broad description of monitoring processes, such as how data will be collected, transferred, stored, managed, safeguarded, and used.
- A basic overview of the anticipated evaluations, including design of and timing for the baseline study, midterm evaluation, and interim/final evaluation. Sections 1.3.6 and 4.4 provide detailed information on the interim/final evaluation.
- A detailed description of M&E staffing; and a basic plan for building capacity of all partner staff who will participate in any form of data collection, analysis, or use. (It is not necessary to provide a complete M&E Staffing and Capacity Development Strategy at the application stage, although this will be required as part of the full M&E Plan submission post-award. Section 3.7 provides detailed information on what should be included in the M&E Staffing and Capacity Development Strategy.)
- A complete, itemized budget for M&E, which should include key monitoring staff, data collection, MIS development and management, logistics, key evaluation, software, hardware, supplies, capacity building, the midterm evaluation, and any other costs related to M&E. The budget should *not* include studies or formative research conducted as part of R&I phase; it should not include costs for baseline, final, or interim evaluation. (Note that midterm evaluations typically cost between $200,000 and $300,000.) BHA expects that applicants will allocate 3%–5% of the total activity budget for program monitoring and the midterm evaluation. The total activity budget comprises the sum of all funds that will be applied to any part of the proposed activity, i.e., funds from Title II (for food commodities; Internal Transportation, Storage and Handling; and Section 202(e)), Community Development Funds, USAID Missions, the partner, and other USAID and non-USAID sources.

Table 1. M&E Plan Submissions and Revisions through the Award Cycle

Component	Application	Start-Up and Refinement Phase	Annually	Midterm	Interim/Final Evaluation
Theory of Change (TOC)	Diagrams should identify how interventions will produce Outputs, Outcomes, and Goal. Complementary information in narrative or table outlines evidence base, clarifies rationales, assumptions, and associated risks.	Refine with additional evidence and improved knowledge through the refinement phase.	Revise and resubmit annually 2 months prior to PREP submission.	TOC should be updated based on findings from the Midterm Evaluation (MTE).	Interim/final evaluation tests TOC (intermediate outcomes only)
Logical Framework (LogFrame)	Must correspond to TOC but result statements should only include those within manageable interest of activity. Must include at least one indicator for each result; and targets for BL indicators for Purposes and Sub-Purposes.	Refine with additional evidence and improved knowledge through the refinement phase.	Revise and resubmit annually 2 months prior to PREP submission.		
Indicator Performance Tracking Table (IPTT)		All submissions must have targets for every BL indicator and required disaggregates. After baseline, the baseline values must be included. All submissions must include base values, annual targets, and LOA values for all monitoring indicators. Base values and disaggregate targets must be included by end of refinement stage.	Revise and resubmit twice annually: with ARR and 2 months prior to PREP submission. For ARR: include reporting year actual values; revised future targets; and any changes to indicators.		Add endline and LOA actual values.
Reference Sheets (PIRS and QuIPS)		Submit PIRS for custom BL indicators 2 weeks before baseline planning workshop or four weeks before the baseline survey training. Submit PIRSs and QuIPS with every M&E Plan submission (2 months before PREP submission).	Revise and resubmit 2 months prior to PREP submission.		
Monitoring Strategy	Describe planned data collection, transmission, storage, management, safeguarding, and utilization, including description of monitoring databases. Include brief description of the plan for MIS, software, and hardware.	Refine and expand through the refinement phase. Include: qualitative and quantitative data collection methods and instruments; data flow diagram/matrix; data quality assurance strategy; data management and safeguard plan; and data utilization strategy.	Revise and resubmit annually 2 months prior to PREP submission.	MTE will review monitoring processes. Strategy should be updated after MTE.	

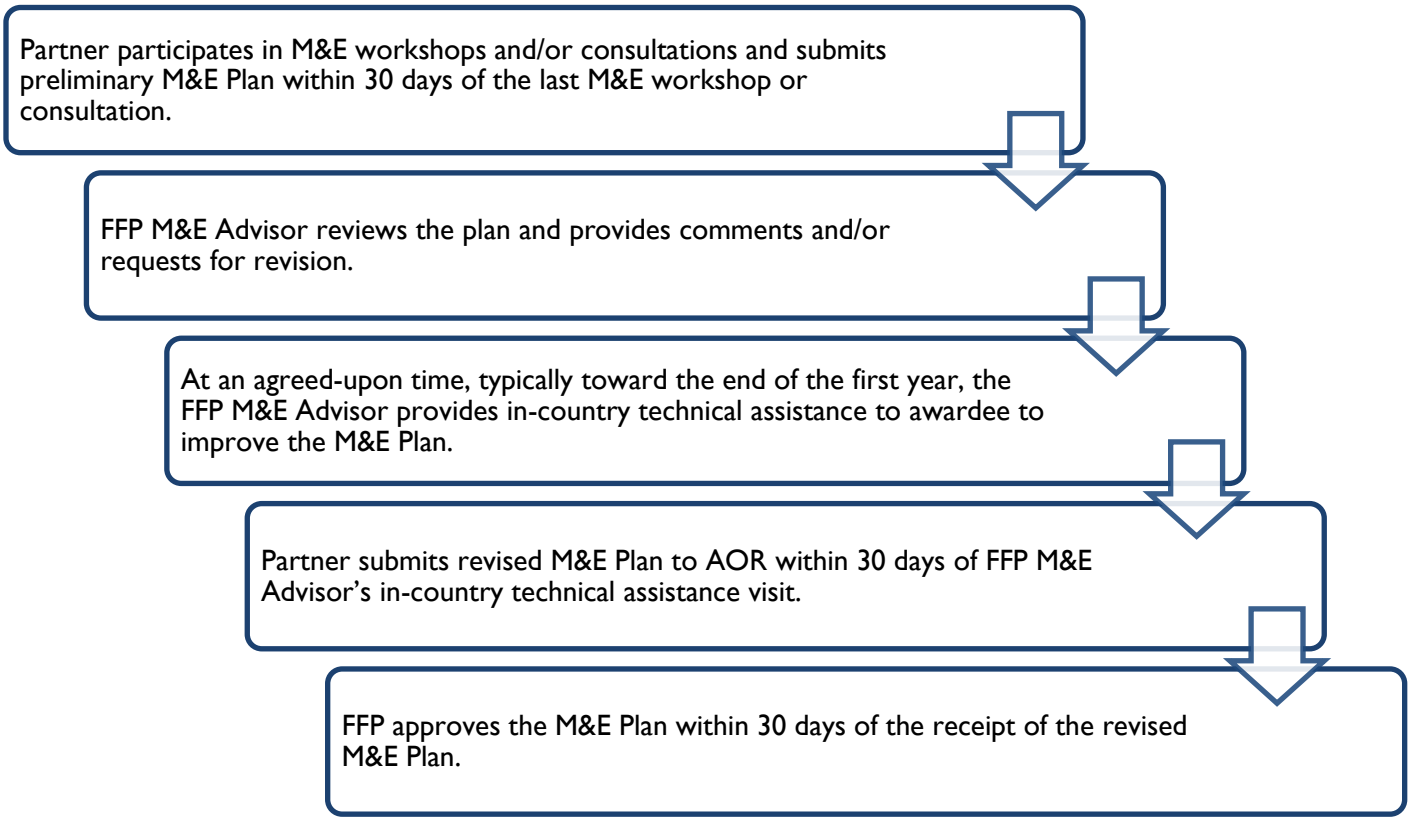
Component	Application	Start-Up and Refinement Phase	Annually	Midterm	Interim/Final Eval.
M&E Staffing and Capacity Development Strategy	Personnel, qualifications, numbers of staff, and roles in Monitoring Strategy and other M&E activities. A brief description of the strategy to strengthen the technical capacity of the M&E team including M&E staff from all partners.	Beginning at M&E workshops, refine and detail to correspond to Monitoring Strategy and other M&E activities through the refinement phase.	Revise and resubmit annually 2 months prior to PREP submission. In ARR, describe achievements of capacity building during reporting year.		
Baseline Study	Basic description of baseline study plans, including the anticipated timing. The RFA will indicate whether BHA plans to design the evaluation using experimental methods. For such a design, the partner is expected to collaborate with the evaluation partner to randomly assign the target (treatment) communities and control communities. The description should highlight partners intent to collaborate.	Prior to baseline planning workshop, or four weeks prior to the baseline survey training, partners identify custom indicators to be collected in baseline and endline surveys and submit PIRS for each.	Enter baseline study report, and values and endline targets in IPTT, with ARR submission (after baseline is completed).		
Midterm evaluation (MTE)	• At award, submit basic description of MTE plans, with proposed timing and description of contracting (partner may contract an external team to complete MTE and/or collaborate with BHA-contracted evaluation team). • Submit complete statement of work (SOW) for AOR approval within 15 months of award or as indicated in the award language. SOW should clearly identify the timeline for conducting the MTE. • For partner-managed contracted MTE's, partner must submit final report within 36 months of award; upload final report to DEC; upload final report with next FY ARR; and submit data to DDL within 30 days of AOR approval. • Submit a Post-MTE Utilization and Action Plan within 45 days of final approval of MTE report.				
Implementation Quality Review (IQR)	BHA and the partner will jointly decide whether an IQR will be conducted in lieu of or in addition to a midterm evaluation.				
Interim/Final Evaluation	• At application stage: provide basic description and timing (i.e., during Year 4 at the same month the baseline will be conducted). • After baseline, enter endline targets for BL indicators into IPTT and ARR. • After interim or final evaluation, enter endline actual values for BL indicators into IPTT and ARR. • If a cost extension is granted, provide updated final evaluation targets in the Year 5 IPTT. Provide final evaluation actual values after final evaluation values are available (in final IPTT and ARR).				
M&E Plan for the Extension Phase	Based on performance, should BHA decide to extend the life of the award, the partner is expected to develop an M&E plan for the extension phase. This plan will be developed in consultation with the AOR and BHA M&E team.				

1.3.2 M&E REQUIREMENTS DURING THE START-UP AND REFINEMENT PHASE

M&E Plan Submission

Post-award, partners are expected to gradually build on the abridged M&E Plan that was submitted at application and develop this into a complete M&E Plan that fulfills all requirements outlined in this guide. Partners are required to attend one (or more) BHA M&E workshops and/or consultations, which aim to further clarify BHA's requirements for M&E Plans and build the technical capacity of partner staff. BHA M&E Advisors will provide timely feedback and support as the applicants develop and refine the activity M&E Plan during the first award year and as needed throughout the life of the activity. The process and timeline for submission and approval of the first complete M&E Plan are illustrated in Figure 1.

Figure 1. M&E Plan Submission Timeline



Partner Involvement in the Baseline Study

BHA will manage the baseline study and interim/final evaluation for each resilience activity; and a research/evaluation partner will carry out the data collection and analysis, and provide the data and report to BHA and the activity implementing partner. BHA may use either a pre-post final performance evaluation design or an experimental or quasi-experimental impact evaluation design.

In a pre-post final performance evaluation design, the baseline data will be collected at the population level, i.e., it will include households from all communities where implementation will take place. In this case, the baseline values are generated for population-based outcome and impact indicators that will be compared to population-based values collected during the endline survey (for the interim/final evaluation) in the fourth year of the activity. If a pre-post design is used, the baseline data should be used

by the partner to refine the strategy for direct participant targeting as well as secondary adoption of key outcomes that will lead to population-level changes.

By contrast, in an experimental or quasi-experimental baseline design, the baseline data will likely be collected from both a treatment group (i.e., where the activity will be implemented) and a control group (i.e. where the activity is not being implemented). Alternatively, there may be an alternative design such that all areas receive some type of intervention and each group or cohort (or implementation area) will be compared to each other. If an experimental or quasi-experimental design is used for the baseline study and interim/final evaluation, the baseline data will be used to determine (a) whether there is any difference in key outcome indicator results between the treatment and control/comparison groups; (b) to calculate the power for the endline survey; (c) run statistical analyses; and (d) refine the intervention design. The BHA M&E Advisor will provide specific guidance on how the activity implementing partner can work with BHA and the research/evaluation partner to interpret and use data from the baseline to refine the design of the activity and set targets.

Applicants and partners should refer to Annex I for specific information on the baseline study.

1.3.3 ANNUAL M&E REQUIREMENTS

The M&E Plan is a living document and should be updated throughout the life of an activity. At minimum, partners must submit a complete, updated M&E Plan annually two months prior to the Pipeline and Resources Estimate Proposal (PREP) submission and an updated IPTT with the Annual Results Report (ARR) submission every November.

M&E Plan Submission Prior to PREP

After the initial M&E Plan has been approved, an updated, complete M&E Plan package must be submitted two months prior to the PREP submission each year. The updated plan should reflect all changes to the monitoring system, indicators, targets, staff capacity building strategies, or other aspects of the M&E strategy or system. Adjustments to the TOC (and any other parts of the M&E Plan) based on contextual shifts or adaptations to the program design should also be included.

Annual Results Report Submission

At the beginning of each fiscal year, the partner must submit an updated IPTT as part of the ARR submission. The ARR is an opportunity to request changes to future targets in the IPTT for all monitoring indicators. The reporting FY targets cannot be changed as they serve as points of comparison to the actual values for the FY reporting cycle. For example, for the Fiscal Year (FY) 2021 ARR in November 2021 (FY 2022), a partner may only request changes to FY 2022 and FY 2023 targets. Future annual, LOA, and endline targets may be changed at any time with justification and AOR approval, except in the final year of the award for endline targets.

BHA annually posts guidance on the content and submission schedule for ARRs on the BHA Implementation and Reporting website.

Quarterly Reporting

Partners are required to submit quarterly reports as indicated in the award language. In lieu of a fourth quarterly report each FY, partners submit the ARR following the BHA ARR guidance. Quarterly reports should draw on qualitative and quantitative data collected, identify changes in the operating context from the previous quarter, and identify potential changes for the next quarter. Partners should refer to the award agreement for additional details about quarterly reporting requirements.

1.3.4 MIDTERM EVALUATION

BHA requires midterm evaluations (MTE) to be carried out for all resilience awards that are four or more years in duration. Midterm evaluations are process evaluations and explore the quality of implementation including social and behavior change related interventions, how well the activity is following implementation plans and meeting targets; the acceptability of the methods employed to the participant population; and signs of changes that participants associate with activity interventions.

The MTE takes place approximately midway through the LOA. Because the MTE focuses on implementation, the in-country data collection should take place when most interventions are happening; the timing does not need to correspond with that of the baseline or endline surveys.

The partner, M&E Advisor, AOR, and USAID Mission determine the timing of the MTE within the first 12-15 months of the award. Additionally, partners must obtain AOR approval on the following:

- MTE Scope of Work: The first draft should be submitted to the AOR within 15 months of the award or as indicated in the award language.
- Team leader and technical sector experts: These individuals must be identified, and their credentials submitted to the AOR for approval no later than 6 months prior to data collection.
- The MTE protocol: At least one week prior to primary data collection begins.
- Final report: Within 36 months of award.

The MTE may be managed either by the partner or BHA. If managed by the partner, the partner must ensure that the evaluation report is uploaded to the Development Experience Clearinghouse (DEC), and any quantitative datasets to the DDL, within 30 days after BHA approval of the report. (If the MTE is managed by BHA, BHA will ensure that the report and quantitative datasets are uploaded to the DEC and DDL, respectively.) In all cases, the partner must upload the final MTE report to the BHA Award Results Tracking (ART) tool in Abacus, previously the partner reporting tool (PRT), with the ARR for the FY in which the report is approved.

Within 45 days of BHA approval of the final MTE report, the partner, in collaboration with the country BHA Officer, must submit a MTE Utilization and Action Plan to the AOR and USAID Mission for their approval.

Refer to Section 4.2 and Annex II for additional guidance on BHA midterm evaluations and the MTE Utilization and Action Plan.

1.3.5 IMPLEMENTATION QUALITY REVIEW (IQR)

The R&I approach allows partners to use the first phase of the activity (typically 12 months) to conduct formative research/studies, pilot studies, and community consultations to refine the design of the activity. The midterm evaluation is typically conducted midway through the LOA and the report is typically available in Year 3 or 4. Therefore, a partner may not have adequate time to fully implement recommendations from a midterm evaluation. As such, BHA, in consultation with the partner, may decide to conduct an implementation quality review (IQR) in lieu of or in addition to a midterm evaluation. The IQR will have a narrower scope compared to a midterm evaluation; involve shorter data collection time; may use a much smaller team compared to a midterm evaluation; and will not require a detailed report.

The IQR will use rigorous, qualitative methods to collect and analyze primary data collected from the field, and it will also analyze existing activity monitoring data (both quantitative and qualitative). The review will primarily focus on implementation quality, sustainability, and systems/institutional strengthening. This approach will provide quick feedback and enable the partner to make corrective measures needed to sustain and demonstrate high-quality performance.

The findings of the IQR (and/or the midterm evaluation) in combination with the interim/final evaluation findings, may be used to inform a potential extension decision.

The IQR may be conducted by: (1) an internal team led by the BHA M&E team with participation from BHA technical team members, partner headquarters and/or regionally-based technical specialists, and technical staff from USAID Mission; (2) an external research/evaluation partner; or (3) a hybrid approach including both external consultants and USAID and partner technical staff.

1.3.6 INTERIM/FINAL EVALUATION

A final evaluation is required for all BHA resilience activities to assess performance against stated objectives and approved targets, except in rare circumstances when BHA may waive the requirement for a final evaluation. The interim evaluation will be managed by BHA and conducted by a research/evaluation partner.

For R&I activities, BHA will manage an interim evaluation that will be conducted in the fourth year to help inform a possible cost extension of the activity. If the activity is not extended, the interim evaluation will serve as the final evaluation. In the event that an activity receives a performance-based cost extension, BHA will conduct a final evaluation at the end of the extension phase.

The interim/final evaluation will use a mixed method approach and will include both a quantitative and qualitative component. The qualitative component will address sustainability; local systems and institutional capacity strengthening activities; gender equity in decision making; social accountability and governance; and other relevant topics.

If a pre-post final performance evaluation design was used at baseline, then the interim/final evaluation will collect endline values using the same population-level sampling frame as the baseline.

If an experimental or quasi-experimental design was used to establish baseline values, the baseline and endline surveys may be either population- or participant-based. As such, the interim/final evaluation will use a different quantitative approach, such as comparing treatment groups (i.e., in areas where the activity has been implemented) with control groups (i.e., in areas where the activity was not implemented).

Regardless of the quantitative component design, the interim/final evaluation will focus on examining progress achieved toward **intermediate outcomes** (e.g., practices and behaviors promoted by the activity) and progress achieved toward long-term sustainability, institutional strengthening, gender equity, social accountability, governance, and other cross-cutting outcomes. In other words, **the interim/final evaluation will primarily focus on outcomes that are directly attributable to well-designed, well-implemented activity interventions** (rather than high-level outcomes such as prevalence of poverty or stunting which may only be partially attributable to successful implementation and sound activity design).

Partners must upload the final report and record endline indicator values as part of the ARR data entry for the FY during which the endline survey was completed.

Chapter 4 and Annex III provides additional information on the methodology and scope of interim/final evaluations that use the pre-post evaluation design. BHA will provide applicants and/or partners with additional information on the methodology and scope of impact evaluation designs on a case by case basis, when appropriate.

1.4 SUBMISSION REQUIREMENTS FOR REPORTS AND DATASETS

In compliance with the ADS and award requirements, all USAID partners must upload all data and reports generated with USAID funding to the appropriate USAID websites. Reports should be Section 508-compliant and uploaded to the Development Experience Clearinghouse (DEC); machine-readable datasets and accompanying documents (e.g., codebooks) should be uploaded to the Development Data Library (DDL); and required data should be entered directly into ART Partners should refer to the award for specific requirements for submission of data and reports.

Partners should upload the final report to ART as part of the ARR for the FY in which the report was finalized. If a final report is not available at the time of ARR submission, partners should note this in the ARR narrative and request to upload it later.

If the partner or BHA has any concerns that publicizing any data or documents could cause harm to activity participants or to the partner, the partner should seek AOR guidance and/or approval for not making the report or data publicly available.

Submitting Reports to the Development Experience Clearinghouse

Partners (or contractors, in the case of externally contracted evaluations) typically submit reports to the DEC within 30 days of AOR approval or as indicated in the award language. Partners should refer to the award for award-specific requirements.

When submitting to the DEC:

- Select the appropriate document type, e.g., "Final Evaluation Report" for final evaluation reports, "Other USAID Evaluation" for MTE reports, and "Other USAID Supported Study/Document" for baseline study reports, qualitative study reports, and ARRs.
- Select "Food Aid Programs" as the primary subject.
- Include all appropriate USAID thesaurus terms in "Additional Information": "Title II Non-Emergency," "Food Security," "Malnutrition," "Child Nutrition," "Maternal Nutrition," "Agriculture," and "Development Assistance."
- Depending on the activity's interventions, also add appropriate thesaurus terms to "Additional Information": "Maternal and Child Feeding Programs," "Maternal and Child Health Care," "Sustainable Agriculture," "Livelihoods," "Family Planning," "Orphan and Vulnerable Children," "Sanitation," "Hygiene," "Humanitarian Assistance," "Disaster Recovery," "Disaster Relief and Response," "Displaced Persons," "Natural Resource Management," "Vulnerable Groups," and "Poverty Reduction."
- Upon submission, send the link of the uploaded document(s) to the AOR and the responsible BHA M&E team member.

For more information on DEC submission requirements, please refer to ADS 540: USAID Development Experience information and the Development Experience Clearinghouse website.

Submitting Data to BHA and Development Data Library (DDL)

Partners and contractors must upload data and supporting documents (e.g., codebooks) from surveys, studies, and evaluations that were conducted by the respective partner or contractor. All data must be in machine-readable format. These documents must be submitted electronically to the AOR by email or through a secure website within 30 days of BHA's approval of the final report. With the approval from the AOR, the dataset and supporting documents must be uploaded to the Development Data Library

(DDL) in accordance with USAID ADS 579 and 579maa: USAID Development Data and the award standard provisions.

Please note that because USAID will make the data available to the public, **the datasets should not contain any personally identifiable information (PII)** that would enable users to identify any survey respondent.

In order to submit data to DDL, the partner must: 1) create a data asset, (2) provide the metadata, which include supporting documents and additional information to help potential users determine the relevance of the dataset to their particular interests, then (3) submit the datasets to DDL, in accordance with the steps and requirements articulated on the DDL website.

BHA will submit the approved machine-readable quantitative datasets and supporting documents to DDL for baseline studies and interim/final evaluations.

Resources

- USAID ADS 540: USAID Development Experience provides policy directives, required procedures, and roles and responsibilities governing the submission of materials to the DEC.
- USAID website is the largest online resource for USAID-funded technical and activity materials.
- USAID Development Data Library is a public repository of USAID-funded, machine-readable data.
- USAID ADS 579: USAID Development Data provides guidance for complying with the requirement for all quantitative data to be stored in a central database (Frequently asked questions: USAID Open Data Policy).
- USAID ADS 201.3.5.8 provides information on monitoring data quality.
- Project Open Data's Common Core Metadata Schema v1.1 provides guidance to support the use of common core metadata to list data sets.

CHAPTER 2. THEORY OF CHANGE (TOC), LOGFRAME, INDICATOR PERFORMANCE TRACKING TABLE (IPTT), PERFORMANCE INDICATOR REFERENCE SHEET (PIRS) AND QUALITATIVE INQUIRY PLANNING SHEET (QUIPS)

2.1 THEORY OF CHANGE

2.1.1 OVERVIEW

Every BHA award application must include the TOC to achieve the intended goal. The TOC describes how all the Outputs from the proposed interventions and Outputs from interventions from other actors/stakeholders will contribute to the Outcomes and, ultimately, the activity Goal. The TOC must be developed based on evidence and should include two parts: (1) a diagram or set of diagrams that illustrates the pathways of change that connect the activity results statements and, (2) complementary information presented either in a narrative or table that further explains and adds to the plausibility of the pathways depicted in the diagram(s).

The TOC diagrams should show not only how Outcomes will be achieved but also how key Outcomes will be sustained over time. For example, TOC diagrams that include provision of extension services to farmers by activity staff should demonstrate how other service providers will support farmers to continue implementing key agricultural practices beyond the life of the activity.

BHA expects changes in high-level Outcomes to be measurable at the population level in the final evaluation. Therefore, in addition to interventions that directly influence the practices and behaviors of participants, the TOC should also depict interventions that will contribute to the whole population. For example, the activity may prepare training participants to act as change agents in their communities and/or broadcast promotional messages about key practices/behaviors through local radio programs.

The TOC must also show contextual factors that are not addressed directly by the activity's interventions but that are necessary for hypothesized changes to occur. This includes anticipated Outputs and Outcomes from concurrent interventions (e.g., other donor-funded activities or host government programs), as well as contextual conditions necessary for the desired changes to occur.

The TOC should be modified throughout the activity and revised TOC must be submitted with the annual M&E Plan submission two months prior to PREP submission. At a minimum, the TOC should be reviewed by program staff at minimum annually, and whenever there is new evidence or when there are changes in the context that affect assumptions or hypothesized pathways of change. When submitting a revised TOC, the partner should describe and justify all modifications and explain how changes to the TOC will affect activity design and implementation. All changes to the TOC must also be reflected in the other required M&E documents (particularly the LogFrame, IPTT, PIRS, and QuIPS).

Box 1. Definitions of TOC Terms

Pathways: The sequence in which Outputs and Outcomes are expected to lead to the Goal. In the TOC diagrams, pathways are depicted using arrows.

Preconditions: Preconditions are the conditions, Outputs and Outcomes that must exist before a higher-level/later Outcome can be achieved. For example, preconditions might include infrastructure improvements; Outcomes stemming from use of the infrastructure; policy changes; or conditions of the political, social, cultural, or natural environment.

Outcomes: Outcomes are changes that are expected to happen when all necessary and sufficient preconditions are met. These might include changes in the ecologic, economic, or governance environment; people's knowledge or attitudes; or cultural beliefs or practices.

Outputs: Outputs are immediate products of interventions, such as trainings delivered, goods or services provided, learning or advocacy events held, or radio communications broadcast. Outputs are typically preconditions for higher-level Outcomes.

Results Statements: Parts of the TOC that are within the manageable interest of the partner (and will be carried over into the LogFrame).

TOC Diagrams: The TOC contains a set of diagrams that use shapes, text, color, and arrows to show the hypothesized pathways of change from Outputs to Goal, plus the critical rationales and assumptions underlying the TOC.

Rationales: The underlying logic and evidence that support the plausibility of connections in a pathway that may not be obvious to the reader. This includes facts or other information to explain why a precondition is necessary and sufficient to ensure an Outcome. For example, a rationale might refer to literature or other evidence that shows that infants in high mental stimulation environments have better growth outcomes compared to those who are in low-stimulation environments. Full explanations of rationales are usually included in the TOC complementary documents.

Assumptions: An assumption describes the contextual or environmental factors or conditions that are out of the control of the activity but have significant influence over the success of the TOC overall or some portion of it. For example, a common assumption for the achievement of a TOC

2.1.2 DEVELOPING THE THEORY OF CHANGE

The process of defining a TOC should start from the activity Goal and work down through the various pathways. Immediate preconditions for the highest-level Purposes and Outcomes in a TOC are typically not direct Outputs of activity interventions or existing conditions. Therefore, other preconditions are necessary before the immediate preconditions can be achieved. These lower-level preconditions must be portrayed in the TOC. The process of identifying preconditions should continue backward, step by step, until reaching an activity Output or a precondition that is outside the control of the activity.

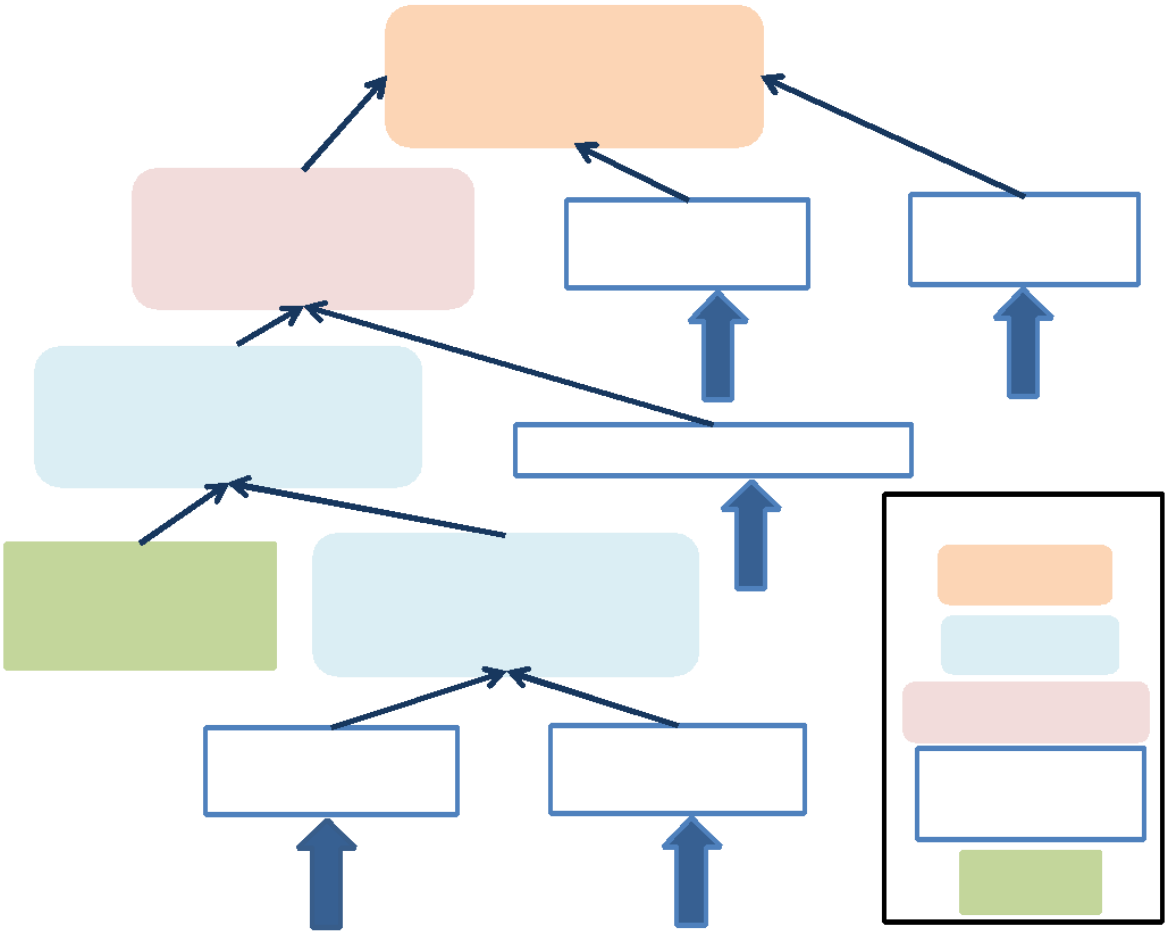
To illustrate, a portion of a TOC is portrayed in Figure 2. This illustration shows how "Household access to nutritious food improved" is an activity Purpose, and a Sub-Purpose is "Household income increased." In this example, one way the activity expects to increase household income is by promoting household members' participation in new income-generating activities (IGAs). To promote participation,

activity staff will train women in IGA skills, and other interventions will address aspects of the economic environment that limit the participation of individuals in IGA skills.

Figure 2 shows that "Household income increased" is a precondition for the Purpose. "Household members participate in new IGAs" is a precondition for "Household income increased." Participation, in turn, is an Outcome of the two preconditions: "Women trained in IGA skills" and "Economic environment for IGAs improved." Women's training is a direct Output of the activity's interventions, and there is no precondition leading to that box. Outputs typically begin pathways. Only an Outcome or condition outside the control of the activity may be a precondition for an Output.

In TOC diagrams, BHA recommends the use of colors, numbers, fonts, shading, and/or shapes to distinguish the different elements of the TOC and the LogFrame and to distinguish Outcomes from preconditions outside the activity's control from those influenced by the activity (see Figure 3 as an example). Every TOC diagram should include a key that explains the significance of the colors, fonts, shading, etc.

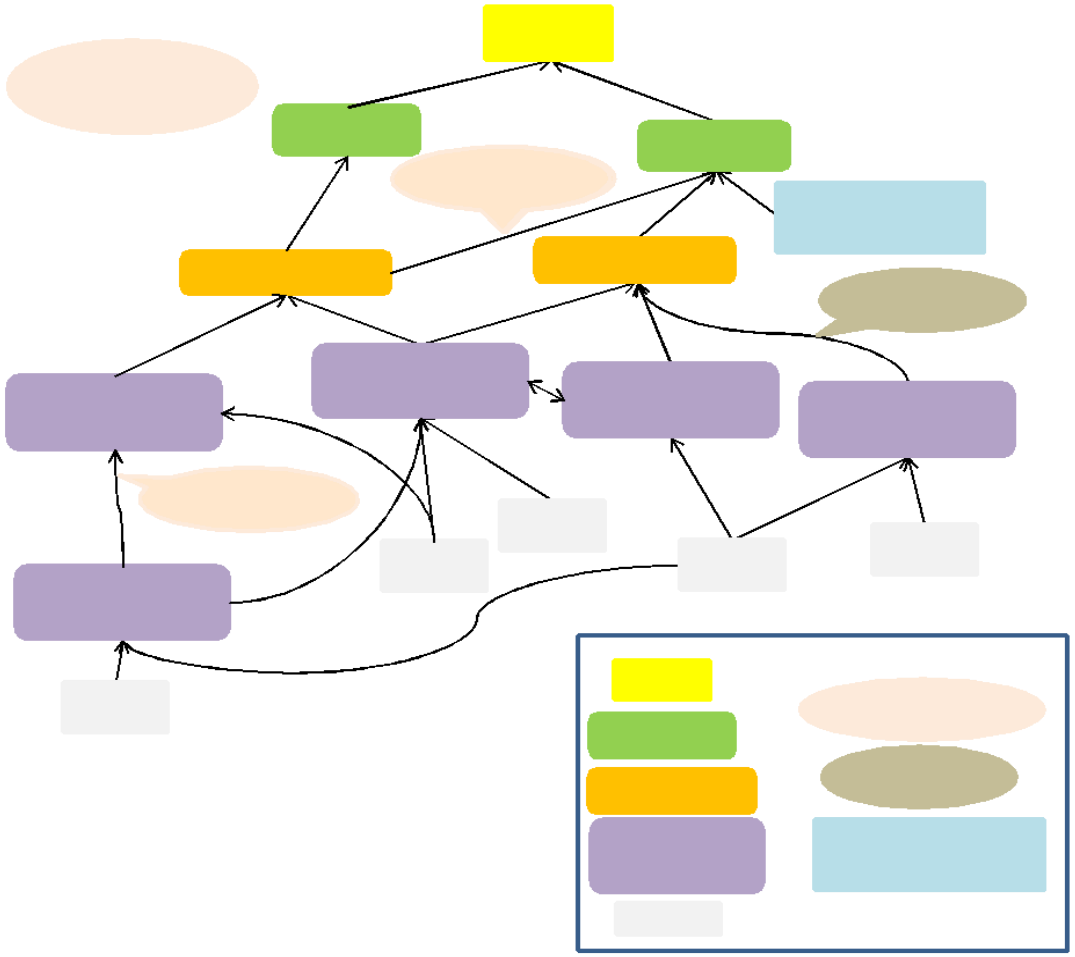
Figure 2. Example of Preconditions/Outcomes in a TOC



Assumptions and rationales should be inserted in TOC diagrams using a unique shape or color. An assumption or rationale underlying or justifying a pathway between two Outcomes should be shown in a shape that points toward the arrow that connects those Outcomes. If the text needed to describe an assumption or rationale makes a diagram too crowded, a simple identifier (for example, A1, R2, etc.) that refers to a description of the assumption or rationale in the TOC narrative or an attached table may be positioned on the diagram to show its relationship to the related Outcome, Output, or pathway (see Figure 2).

All preconditions, assumptions, and Outcomes in the TOC should be stated as results, not processes. For example, the results of training could be "Caretakers of children aged 6–36 months trained to prepare foods to complement breastfeeding," not "Activity trains caretakers of children aged 6–36 months to prepare foods to complement breastfeeding."

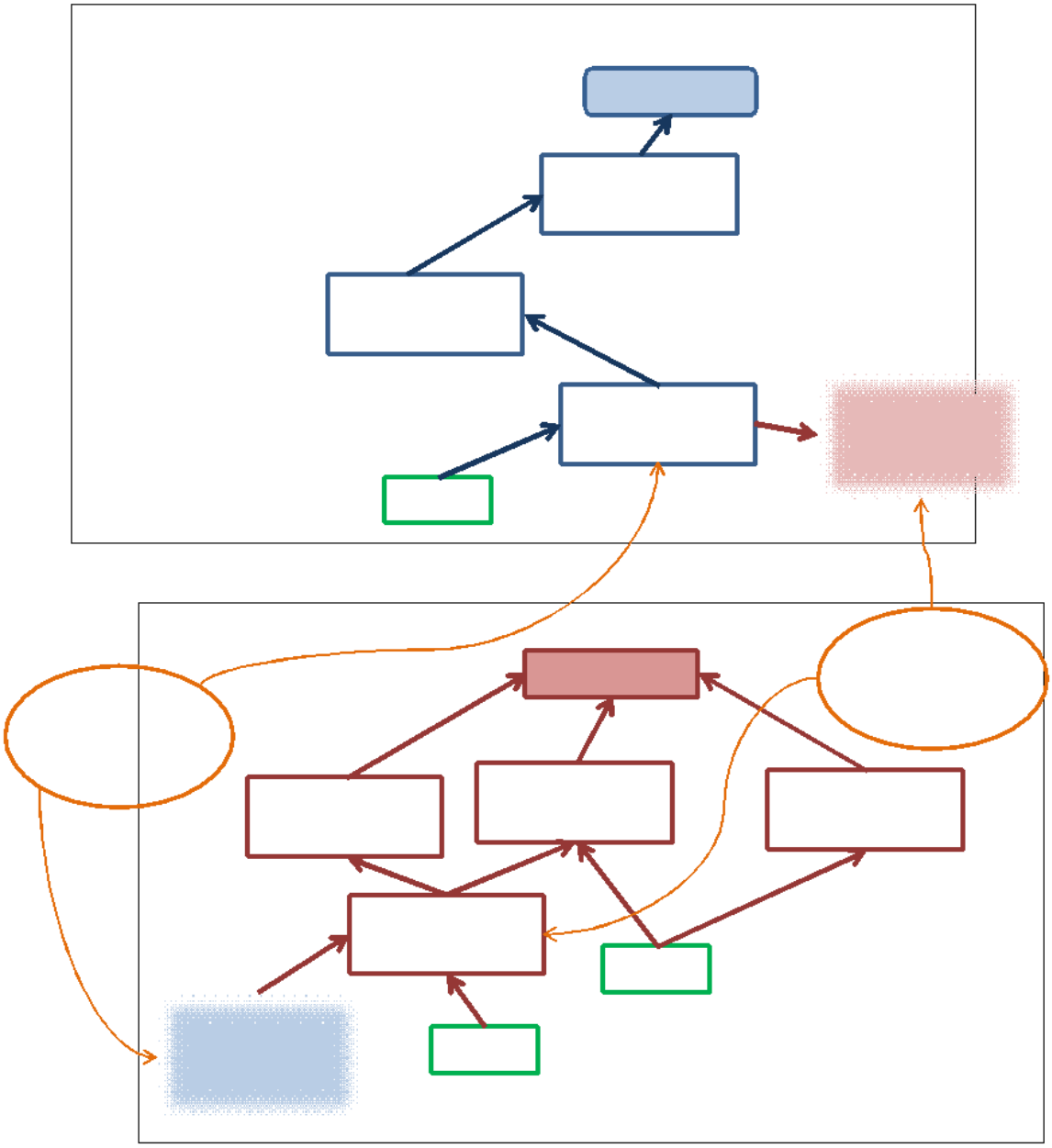
Figure 3. Example Showing Color and Shape to Distinguish TOC Elements



Because BHA resilience activities are multi-sectoral and complex, an activity TOC diagram may not fit onto a single page. To keep the diagrams reader-friendly while maintaining detail, the TOC may comprise a set of diagrams (e.g., one summary diagram with each purpose and additional purpose-specific diagrams). Some pathways in the TOC may overlap across Purposes or Sub-Purposes. Figure 4 shows one way that a crossing might be depicted by repeating the Outcome that appears on the page with the pathways leading up to it on a second page where the Outcome is a precondition for a higher-level Outcome. Likewise, the higher-level Outcome on the second page could be repeated on the first page. Color, shape, or formatting should be used to identify a precondition/Outcome that also appears on another page. (Note that the orange ovals and arrows are not part of the TOC diagram. They are added in this figure only to direct the reader to the repeated elements.)

BHA requests that, in addition to the set of Purpose- or Sub-Purpose-level diagrams, the partner include a single-page, all-inclusive diagram. BHA does not require any specific software for developing and presenting TOC diagrams.

Figure 4. Showing Pathways That Cross Diagrams in a TOC



Articulating Pathways in the TOC

When designing an activity, it is important to analyze each pathway to determine how critical the pathway is to achieve the purpose. In analyzing the pathway, ask: *Will the activity fail to achieve the purpose if the outcome (sub purpose) is not achieved?* There are interventions that contribute to the improvement of the purpose level outcome, but they are not the drivers of the outcome for a particular context. For example, there could be multiple pathways to increase access to and consumption of nutritious food: investing in home garden; improving knowledge on nutritious food; investing in skill building, investing in livelihoods development so people can buy nutritious food; investing in poultry rearing; and other pathways.

In order to increase access to and consumption of nutrition food, the applicant may want to intervene in several pathways depending on the following factors: (1) the potential contribution of each strategy to achieve the outcome of interest; (2) the effort required to implement those strategies; (3) the asset base and livelihoods of the target groups; (4) the potential risks for the partner and/or community members; and (5) and the operational and cultural context.

Once a pathway (or set of pathways) is determined, analyze each intermediate outcome that would contribute to the sub-purpose. The interventions that are proven to be most critical and effective in achieving the sub-purpose should appear in the TOC (and later be transferred to the LogFrame). In analyzing the sub-purpose pathway, ask the following questions:

- What are the key drivers of the sub purpose?
- How likely will this intervention or set of interventions influence the outcome in this particular context?
- Is this intervention proven to be effective by independent reliable evaluation or research?
- How critical is this intervention to achieving the sub purpose (i.e., what happens if the intervention fails?)?

Through these questions only the most critical and effective drivers/influencers will be identified; interventions that only marginally contribute to the sub purpose will be dropped. While their inclusion may prove slightly beneficial to achieving results, ultimately these interventions are costly and not essential to activity outcomes. In finalizing a pathway or a set of interventions, an activity design team must analyze the cost and potential benefit of the pathway or the intervention package. For example, an activity may plan to promote keyhole gardens to improve household diet quality. To promote high quality keyhole gardens, an activity needs to provide training on keyhole gardens, vegetable production, crop management, and pest management. In addition, often an activity provides seed. Even though it may appear that the per household cost to promote keyhole gardens is fairly low, the cost for the intervention package for the life of the activity may not appear to be low when considering the related interventions and existing circumstances. In addition, the staff or promoter must take the time out from another intervention to provide the training. In spite of the support, keyhole gardens may not be successful because a household needs access to water, and small livestock and poultry often eat the leaves of new growth. Ultimately, even if it is successful, a keyhole garden may contribute only marginally to improving diet quality and it may not be cost effective to influence the ultimate outcome of interest in a particular context rather it may take resources away from the interventions that could drive the result.

In the end, only those outcomes that are critical and effective in achieving the sub purpose, and are supported by evidence, should appear in the LogFrame.

Breadth, Depth, and Level of Detail in TOC Diagrams

The "breadth" of the TOC relates to the degree to which it includes external influencing factors. "Depth" refers to how far back the TOC goes from the Goal. The TOC's "level of detail" depends on the magnitude of the step between adjacent preconditions and Outcomes in the various pathways of change. The breadth of the TOC submitted with an application should demonstrate comprehensive understanding of the proposed implementation context and be supported by sufficient evidence. The TOC should include the external preconditions that are most likely to affect the Outcomes necessary to the achievement of the activity's Goal, either positively or negatively.

All BHA multi-sectoral food security resilience programs aim to increase target households' access to food. Increasing access to food for extremely poor and vulnerable populations would require increasing income and/or increasing production. Extremely poor and vulnerable households must rely on diverse

income and production sources which depend on seasonality, asset base, capacities, composition of the household, shock exposure, and other factors. For example, extremely poor households may not have adequate and regular access to productive land on which to farm. These households may not be able to meet their dietary needs through farming and may instead need to engage in other income-generating activities or labor to secure food.

While households may be involved in multiple income generating strategies, not all of these strategies afford equal potential to make a sustainable, transformative change to the households' income or access to food. Therefore, the applicant must develop the income pathways based on an economic analysis to identify the potential income strategy or strategies that will likely transform the extremely poor and vulnerable households' income. In addition, not all extremely poor and vulnerable households have similar assets, capacities, opportunities, and vulnerabilities; therefore, the income pathways for different groups of households must be different. The TOC should include differentiated income pathways based on the households' capacities, opportunities, and vulnerabilities as well as based on an economic analysis to identify the pathway that has substantial growth potential and likely to transform the households' access to income or food.

BHA requires greater depth for the pathways directly affected by activity interventions than for those outside the control of the activity. The Output from every activity intervention must be depicted in the TOC diagram, and the pathways must show how every Output eventually connects to the Goal. For preconditions outside the control of the activity, the TOC diagram should show only the highest-level precondition(s) and identify the actors or environmental circumstances that contribute to the precondition, identify sources of risk to the existence of that necessary precondition, and describe the consequences to higher-level Outcomes if the precondition does not exist.

The TOC submitted at application should include enough detail so a reader who is unfamiliar with the activity can understand how changes are expected to unfold. Partners are required to revise the ToC during the refinement phase and submit the final updated TOC with their comprehensive M&E Plan two months after the end of the refinement phase. After the refinement year, the partner then submits the updated M&E Plan annually two months prior to PREP submission.

TOC Complementary Documentation

The TOC complementary documentation **should not summarize or reiterate what is obvious from the TOC diagrams**. Rather, it should add supplemental information that is not easily communicated graphically or expressed in a few words on a diagram. It should refer to evidence used to develop the TOC and help an external reader understand the degree of certainty that the pathways portrayed in the diagrams will occur. The complementary documentation may:

- Add detail about assumptions: for example, describe trends that indicate growing or declining stability and the sources of risk. It can also highlight the Outcomes that are at greater or lesser risk if conditions change or an assumption fails. In cases of instability or high risk, the narrative may be used to describe how the activity will monitor conditions and act to mitigate the risks or effects of the changes.

- Provide text, web links, or references to scholarly or grey literature that justify connections or causal pathways that are not widely known. For example, this may include studies that show children are more likely to be fed a more nutritious diet if both parents contribute to decisions about feeding as opposed to one parent.

- Identify actors outside the activity who are intervening or will intervene to produce Outcomes or Outputs that are preconditions in a TOC pathway; the scale of their intervention relative to

the activity's coverage; the likelihood that the preconditions will be achieved by the time they are necessary for the TOC; and the risks if they are not. It may also describe anticipated collaboration with each actor, how that collaboration will better ensure the preconditions, and how Outputs and Outcomes will be monitored.

- Explain how an intervention with a limited number of participants will result in population-level change. Some activities use community-based interventions whereas others target a large proportion of the population in each community to reach a critical mass, develop activities and processes in such a way that they will self-replicate, or use the activity participants as change agents in the community. The complementary documentation can be used to provide an explanation of the implementation approach if it is not readily apparent in the TOC diagram.

Cross-Cutting Technical Areas

Activities integrate cross-cutting technical areas of gender, environment, community participation, sustainability, and conflict-sensitivity ("do no harm") in different ways. The TOC diagram and complementary documents should clearly show how the activity has integrated these cross-cutting areas.

Some activities implement specific interventions in cross-cutting technical areas to promote attitude changes or practices of a specific group of participants, or to make structural or organizational changes in a community or natural environment. For example, an activity may undertake interventions that transmit key messages related to gender equity to the general community (e.g., through mass media) or at sites of community events (e.g., performing dramas at sites of food distribution and seed fairs). Some work with target communities on kitchen performance tests[2] for better fuel-efficient cooking stoves during nutrition interventions to improve indoor air quality while messaging on more-effective timber harvesting and charcoal production. Others may facilitate discussions about principles of conflict avoidance and resolution among community leaders, married couples, community organizations, or the general community. Pathways depicting these kinds of interventions that directly relate to a cross-cutting Purpose may be depicted in a diagram on a page dedicated to that cross-cutting Purpose, in which case the diagram should show connections to diagrams for other activity Purposes (see Figure 4).

Some activities integrate cross-cutting technical areas into the implementation methods for interventions that contribute more directly to other Purposes and sectors. For example, interventions to promote savings and loans are gender-responsive when their objectives and training consider that men and women might have different objectives for saving or borrowing or face different barriers to saving. In another example, an activity may drill wells for irrigation and foster a municipal governance structure to ensure that this valuable resource, which is increasingly threatened by climate change,[3] is used in a productive and sustainable manner. Many improved agricultural practices benefit the natural environment as well as crop production. In these cases, the cross-cutting integration may be represented in the TOC in the wording of the Outcomes and Outputs in the diagrams for the other Purposes, for example, by using keywords like "men/women," "climate change sensitive," "good environmental practices," "gender-responsive," "gender-equitable," "with broad community input," "inclusive," and "conflict-sensitive."

[2] See: USAID. 2010. "Fuel Efficient Stoves Programs in Humanitarian Settings: An Implementer's Toolkit." Available at: http://www.ifrc.org/PageFiles/95759/USAID_FES_Toolkit_July_2010.pdf.

[3] See: British Geological Survey. 2016. "Groundwater resilience to climate change in Africa." Available at: http://www.bgs.ac.uk/gwresilience/.

2.1.5 REVIEWING TOCs

Partners must review and revise the activity-level TOC during the refinement phase and at least annually prior to annual M&E Plan submission. BHA recommends that partners convene activity stakeholders from all sectors to review the activity's implementation, progress, and factors that have affected either the implementation or outcomes, and examine validity of the underlying assumptions in the current context. The TOC review should be informed by a variety of available data sources, and include the review of existing sector specific research or analysis reports, assessment findings, formative research, qualitative inquiries, community consultations, monitoring data, recent evaluations, learning events, stakeholder consultations, partner and program participants' experiences, and secondary or other sources of information. For annual reviews, monitoring and other contextual data should be collected, analyzed, and organized early in the preparatory process so that the review itself can focus on learning from the data and using that learning to inform programmatic and/or operational change.

Other considerations for undertaking a TOC review include thinking through who should lead the process and who should participate. Partners may choose how to best facilitate the review, but the process should emphasize the use of data to review the progress or lack of progress along the pathways. (A lack of progress does not always mean that there was a problem with the theory. For example, there could have been an issue with implementation.) The TOC review also provides an opportunity for staff to assess how useful the M&E data are in measuring changes and issues of interest to the activity.

During the **refinement phase**, the TOC refinement should be ongoing and focus on identifying evidence gaps and logical inconsistencies, reviewing the sustainability pathways, eliminating pathways that may have marginal contribution to achieve the higher-level outcomes, identifying areas that need further analysis to revise the income pathways, validating assumptions, and refining the design of the activity. (Note that some activities may begin some implementation during the "refinement phase" with AOR approval, e.g. to commence implementation of tried-and-true interventions that have a robust evidence base in the operational context.)

In **early implementation years**, the TOC review should take place at least annually (prior to M&E Plan submission) and use monitoring data to review the TOC pathways and progress from Outputs to intermediate Outcomes, as well as early indications of systemic changes. In addition, the review should also look at the Outputs to promote sustainability and indications of early Outcomes.

After the **midterm evaluation**, the TOC should be revised based on the recommendations from the midterm and begin to verify lower level results and their contributions to expected Outcomes. TOC reviews near the end of the life of the award can verify that higher level Outcomes will achieve final evaluation targets.

Additional, sector-specific reviews with all or a subset of activity stakeholders may also be useful. To facilitate the review of TOC, the partner should gather quantitative monitoring indicator data and qualitative information to understand performance indicators, context, and assumptions. Examples of evidence or situational factors that might prompt a special revision to a TOC diagram and/or narrative at any time include:

- An intervention or intermediate Outcome that failed to influence a higher-level Outcome, even if high quality implementation and all other preconditions in the TOC were met (e.g., household income increased and knowledge of child nutrition improved, but child feeding practices did not improve).

- An intervention Output that contributed to an unanticipated outcome or contributed to other outcomes in unexpected ways. For example, new boreholes drilled near a village that were intended to provide water for the community's livestock attracted a nomadic group that forcibly

took control of the wells.

- A set of Outputs were delivered and the indicator target was met; however, the logical anticipated Outcome from the Outputs not yet achieved, and the activity does not have a good understanding of the quality of implementation.

- Activity monitoring or a special study that revealed that the quality of, or efficiency in, reaching Outcomes depends on additional factors not portrayed in the TOC. For example, some participants may have been quicker to adopt key practices than anticipated or compared to other groups; a gender analysis may have found that previously unrecognized cultural factors are in fact barriers to adoption of key practices; or findings from a qualitative study may indicate that an intervention is not leading to intended change and identify potential implementation changes.

- New research external to the activity that supports previously unknown causal pathways or refutes previously accepted pathways portrayed in the TOC. For example, peer-reviewed literature and/or meta-analyses suggest for the first time that a factor prevalent in the activity area can contribute to stunting.

- Significant changes in the political or environmental conditions of the local context.

The TOC should be revised and submitted annually as part of the M&E Plan submission two months before PREP submission. Partners should provide a clear description of how the TOC was changed and what processes were used to revise and update the TOC. In addition to the annual submission prior to the PREP, the partner may request AOR approval of a revised TOC at any time.

Resources

- The Theory of Change Training Curriculum developed by TANGO International and the TOPS Program includes materials to support TOC design and revision. The TOC Checklist is a useful tool to ensure the completeness and quality of TOC diagrams and complementary documentation. A TOC review facilitator guide is also available on the Food Security and Nutrition Network (www.fsnnetwork.org).

- The Center for Theory of Change promotes best practices for the development and implementation of a TOC. It particularly emphasizes its application in the areas of international development and sustainability.

- The Annie E. Casey Foundation has a *Practical Tool for Action, Results and Learning* that helps in the development and application of a TOC.

- A blog in the Stanford Social Innovation Review discusses pitfalls to avoid when designing and applying a TOC.

- The Overseas Development Institute offers short planning tools on problem tree analysis.

- The Evaluation Toolbox provides a how-to guide for problem tree/solution tree analysis.

2.2 THE LOGICAL FRAMEWORK

2.2.1 OVERVIEW

While a TOC describes the many pathways and assumptions to achieve the activity Goal both within and outside the manageable interest of the activity, the LogFrame should present only the Outputs and Outcomes that are (1) within the manageable interest of the activity and (2) crucial for the partner to demonstrate the most important activity results. A LogFrame is the summary of an activity design, therefore it drives the monitoring and evaluation system.

An activity often does not include plans to intervene in all technical sectors necessary to achieve the highest-level food security, nutrition, and/or poverty outcomes. Often there are other entities/actors who implement interventions in the same geographic area that also contribute to achieve high-level activity outcomes. The LogFrame should only comprise those select results from the TOC that are directly within the manageable interest of the partner and that are absolutely necessary for the partner to demonstrate key results along key pathways. What ends up in the LogFrame should be based on a careful review of the various pathways in the TOC and prioritization of interventions that would likely have the greatest influence on food security, considering the comparative advantage of the partner.

Box 2. Types of Result Statements in a TOC and LogFrame

Goal: The highest-level Outcome to which an activity can contribute. Typically, a Goal cannot be fully accomplished during the award period. Factors beyond the control of the activity must also be addressed before the Goal can be fully accomplished, or achievement will take longer than the life of award (LOA). The Goal is the ultimate objective of the activity and is directly linked to aBHA Strategic Objective (SO) and/or a USAID country Development Objective. An example of a Goal for a BHA activity could be "Sustainable food security in households of XXX province achieved."

Purpose: A key, high-level Outcome that the activity expects to accomplish during the LOA. A Purpose described a desired change in the condition or status of the population in the target area to which the Outputs and Outcomes of the activity's interventions should contribute significantly. An example of an activity Purpose could be "Nutritional status of women of reproductive age and children under 5 years improved."

Sub-Purpose: An Outcome necessary for a Purpose to be achieved. These often include behavioral and systemic changes, for example, adoption of promoted techniques or behaviors; changes in response time, management systems, natural resource conditions, income, or capacities; or shifts in cultural norms. An example of a Sub-Purpose could be "Increased household income from farm and off-farm sources."

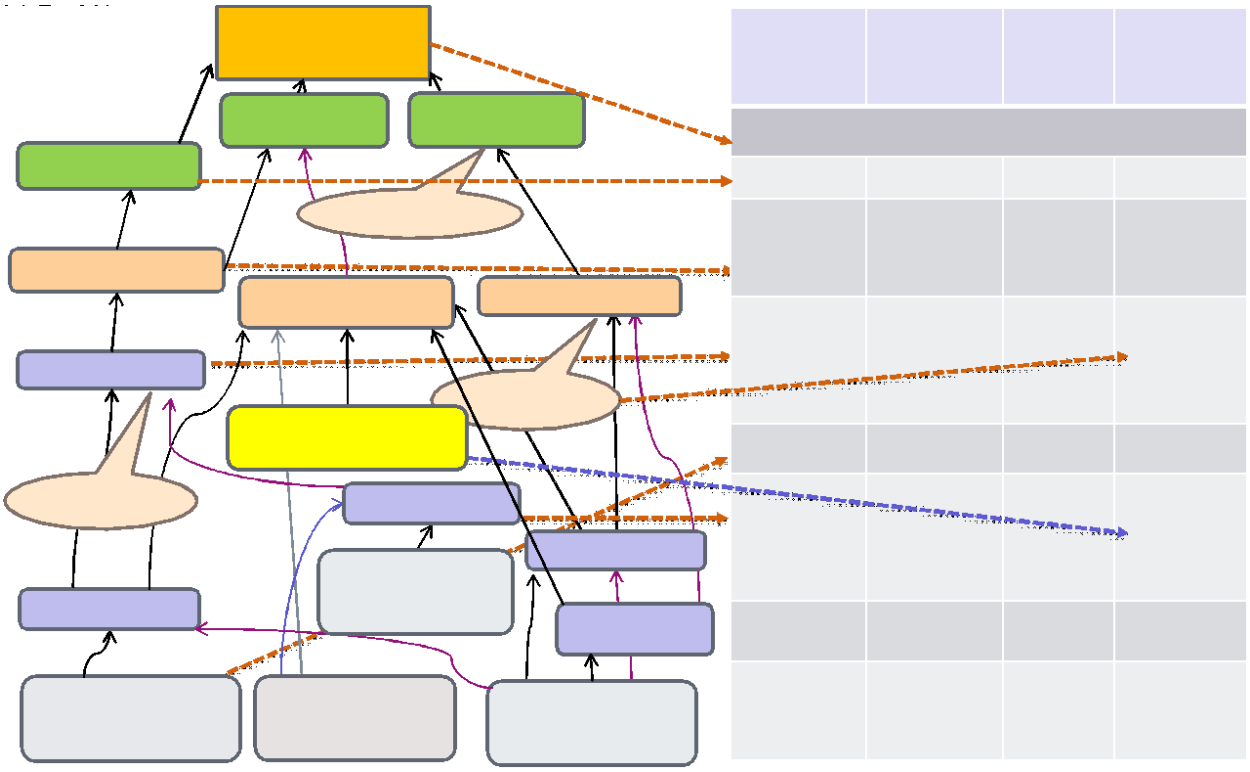
Intermediate Outcome: An Outcome that must occur before a Sub-Purpose or another Intermediate Outcome can be achieved, such as changes in knowledge or attitudes, mastery of skills, and adoption of new methods. Examples include "Increased application of improved farming practices on own land," "Increased consumption of promoted foods," and "Greater participation in growth monitoring." There may be multiple levels of Intermediate Outcomes in sequence along a single pathway.

Output: Tangible, immediate product of an intervention under the activity's control or influence. Examples include "Number of people training," "Quantity of food rations distributed," "Number of groups formed," and "Number of different types of infrastructure rehabilitated or improved."

Note: BHA does not require Inputs in the LogFrame.

A LogFrame uses a standard matrix format to summarize key elements of the TOC in four columns titled "Result Statement," "Indicator/Inquiry," "Data Sources/Collection Methods," and "Assumptions." Assumptions refer to natural, climatic, economic, social, political, environmental, policy, and external programs implemented by other actors. Table 2 provides an example of a portion of a LogFrame, and Figure 5 shows the relationship between TOC and LogFrame components.

Figure 5. Relationships between TOC and LogFrame



BHA requires all BHA required and required if applicable indicators to be included in the LogFrame. It is not necessary to include all Outputs from the TOC in the LogFrame; rather, only indicators that will be used to measure key results (either Output or Outcome indicators) should be included in the LogFrame. For example, stand-alone outputs or deliverables (such as creation of manuals or construction of a warehouse) do not need to be included in the LogFrame as these can be reported on in written form in reporting documents such as the ARR and quarterly performance reports.

Indicator or Inquiry: For each result statement in the LogFrame, there should be at least one quantitative, monitoring indicator or qualitative monitoring inquiry (qualitative monitoring or qualitative study) articulated in the second column.

Quantitative Indicators: There are three broad categories of indicators: *BL indicators* measure conditions at the population level at the beginning and end of the activity; *monitoring indicators* are measured regularly among participants to report annually throughout the LOA; and *contextual indicators* provide information about conditions that are outside the influence of the activity but important to the Outcomes of the activity. All required and required if applicable BHA and Mission indicators must be included in the LogFrame.

At the application stage, BHA requires applicants to propose targets for the final evaluation Purpose and Sub-Purpose indicators[4] in parentheses following the indicator title in the second column. These targets may be presented as:

- A *fixed value*, for an indicator expressed in any unit (e.g., 80% for "Prevalence of exclusive breastfeeding of children under six months of age" or $4 for "Daily per capita expenditures in U.S. Government-assisted areas");
- A *percentage point change* from baseline, for indicators expressed as percentages (e.g., baseline value -10 percentage points for the indicator "Prevalence of stunted children under 5 years of age"); or
- A *percent change* from baseline, for indicators expressed in units other than percent (e.g., baseline value * 130 percent for "Yield of targeted agricultural commodities within target areas")

To establish final evaluation targets at the application stage, partners need to consider the TOC, targets suggested by BHA, proportion of households in target communities are targeted, and existing literature to address information gaps.

Qualitative Inquiry: If the result is best measured through qualitative lines of inquiry, i.e. either qualitative monitoring (QM) or a stand-alone qualitative study (QS), then include a reference to the QM or QS that will be used to monitor progress achieved toward the result in the first column. Include only the primary line of inquiry and indicate the number of the QM or QS (e.g., "QM 1"), using the order in which it appears in the LogFrame to determine the number. In cases where a QM or QS is needed in addition to quantitative indicators, these may also be included in the LogFrame. It is not necessary to include both a quantitative indicator and qualitative inquiry for each result.

Table 2 provides an example of how qualitative inquiries might be included in a LogFrame. The Qualitative Study (QS 1) under Intermediate Outcome 1.1.1 would help to both (1) inform the technical approach; and (2) monitor change in the result statement between the beginning and end of the activity. In this case, the partner is proposing to use the quantitative survey data to help guide the two phases of the qualitative study (e.g., to help select which couples or individuals to interview and understand how many couples are making decisions jointly).

The Qualitative Monitoring (QM 1) under "Intermediate Outcome 1.1.2 Community norms support positive MCHN practices among all families" outlines the qualitative monitoring questions that will need to be answered in order to monitor change toward this result statement. In this case, the result statement (IO 1.1.2) can only be monitored using qualitative methods, so the partner has proposed a monitoring approach that will enable tracking the result without quantitative indicators.

Data source/collection methods: This column should provide a short summary of the data source and data collection method that will be used to measure the indicator. For example, an indicator may be measured by using a questionnaire (data source) which is collected through participant-based survey (data collection method) at baseline and endline or routine monitoring (data collection method). Another example: training events use attendance sheets (data source) and are collected by way of routine monitoring (data collection method).

Be as specific as possible, and avoid generic terms like "project records." This description should provide enough information so that the user of the guide and reviewers would know where to look in the M&E Plan for a more detailed description of the methods that will be used. See Chapter 3, for the various types of data collection methods.

[4] Targets for disaggregates are **not** required in the LogFrame.

Assumptions: Assumptions are factors that are outside the activity's influence that are necessary for an activity Outcome or Output to be achieved. This includes the assumptions in the TOC plus all preconditions that are identified in the TOC as being outside the control of the activity. Interventions implemented by other actors but critical to achieve the food security outcome are also outside of the activity's manageable interest hence should be added to the assumption. Factors that the activity seeks to influence should not be included as an assumption, including the achievement of intermediate Outcomes. For example, for an activity that seeks to change behavior, "Participants are open to changing behavior" may sound like an assumption, but the activity interventions are seeking to change attitudes that would lead to behavior change. Hence, the degree of openness may depend on the quality of behavior change sessions and how well the activity can explain the benefits. Assumptions may be identified and monitored through secondary data and reports, qualitative methods such as interviews and discussions, and quantitative surveys.

Since the LogFrame is organized in table or matrix format, it cannot reflect all pathways in the TOC. For example, an Intermediate Outcome's contribution to more than one Purpose is easily depicted in the TOC diagram using multiple arrows in different pathways. However, in the LogFrame, the same Intermediate Outcome can be included in only a single pathway. For this reason, that outcome should appear in the LogFrame under the Purpose to which that outcome will make the greatest contribution. A LogFrame also cannot reflect the interdependence among Outcomes at the same level. For example, it cannot show that a Purpose of "Improved use of high-quality maternal and child health and nutrition services" is a precondition for another Purpose of "Improved nutritional status of under-5s." The flexibility of the TOC diagram to show interdependencies and multiple pathways is one reason why BHA requires a TOC for every activity.

Table 2. Illustrative LogFrame

Result Statement	Indicator or QM/QS Inquiry	Data Source/ Collection Methods	Assumptions
Goal: Male and female population in Tangail District are food secure			
Purpose 1: Chronic malnutrition in children under 5 years is reduced	BL 5: Prevalence of stunted children under 5 years of age (Baseline -10 percentage points)	Household questionnaire, population-based survey (PBS)	Ministry of Health maintains adequate MCHN health service personnel and supplies in local health centers throughout LOA. Stable health/WASH conditions. There are no widespread disease outbreaks that the health systems cannot address with available supplies, staff, and other resources.
Sub-Purpose 1.1: (MCHN) care practices improved	BL 26: Percent of births receiving at least 4 ANC visits during pregnancy (90%)		
	Custom 1. # of live births receiving at least 4 ANC visits during pregnancy Custom 2. # of women receiving postpartum family planning counseling Custom 3. % of child participants who receive vaccinations on time	Monitoring form/checklist, Routine monitoring	
Intermediate Outcome 1.1.1: Both men and women cooperate to support good MCHN care practices	Custom 4: Percentage of men and women with children under 2 who make maternal health and nutrition decisions jointly QS 1: How does joint decision-making affect support for good MCHN care practices? ● What are the barriers and facilitators to joint decision making? ● Among couples who make decisions jointly, how are MCHN decisions communicated and made? Are these decisions ultimately leading to good MCHN care practices? ● Are there unintended consequences from making joint decisions? ● Among couples who make decisions separately, how are MCHN decisions made?	Household questionnaire, PBS Interviews with couples; data will be collected after Y1 and Y5 data are collected to understand changes in decision-making	
Intermediate Outcome 1.1.2: Community norms support positive MCHN practices among all families	QM 1: What are religious/traditional beliefs and norms among leaders and community members around MCHN? ● How are norms changing around women fasting while pregnant and lactating? ● How are male partners and other HH members providing intellectual and material support to women's nutrition while pregnant? ● How are norms changing around child marriage and early pregnancy (which are known drivers of maternal and child malnutrition)?	Data collection via observation, focus groups; and interviews	There is room within a patriarchal structure that will allow for these changes.
Sub-Purpose 1.2: Household and	BL 27: % of households with access to basic sanitation service (Baseline + 10 percentage points)	Household questionnaire, PBS	Local supplies and prices of soap and

community hygiene improved	BL 17: % of households with soap and water at a handwashing station on premises (Baseline +15 percentage points) Custom 6. % of villages with active sanitation committees supported by village council (100%)		sanitation maintenance materials remain constant relative to incomes
	AM 22: Number of people gaining access to an improved sanitation facility	Monitoring form/checklist	
	Custom 7: % of improved sanitation facilities with feces visibly present on the floor, wall, or area immediately surrounding the facility	Household questionnaire, PaBS	

2.2.2 INDICATORS

The activity indicators in the LogFrame should include:

- All required and applicable BHA indicators;
- All Mission-required indicators (post-award only);
- All custom outcome indicators;
- Key output indicators (i.e., only those outputs that are deemed necessary and important by the partner for tracking performance on key aspects of the activity); and
- Key context indicators (i.e., only those context indicators deemed necessary to monitoring the operational, cultural, security, or other context).

The BHA Indicator List is available on the BHA Food Assistance Implementation and Reporting webpage.

BHA indicators are classified as required (R) or "required if applicable" (RiA). Required indicators are mandatory for all resilience activities. Required if applicable indicators are only required if the indicator is relevant and necessary based on the activity's interventions. Applicability criteria are included in the indicator PIRS and the BHA Indicator List. The "Frequency of Report" column on the BHA Indicator List specifies whether data collection is required annually or only at the beginning and toward the end of the activity. The latter indicators are typically collected at baseline (and during the quantitative survey for the interim/final evaluation), so those are referred to as "BL" indicators.

The BHA Indicator Handbook Parts I and II include PIRSs for BHA indicators, including questionnaires and tabulation instructions. The BHA Food Assistance Implementation and Reporting webpage has links to the handbook.

Activities awarded before FY 2019 may use BHA monitoring (M) indicators that were added to the BHA Indicator List after their award, but partners are not required to do so unless the relevant USAID Mission requests this. For partners with awards dating before FY2019, the indicators collected at baseline should be collected in the final evaluation survey, i.e., new BL indicators do not apply.

The USAID gender policy requires all USAID activities to collect appropriate sex-disaggregated data for all people level indicators, ask clear questions about gender roles to reveal both intended and unintended positive or negative changes, and develop indicators designed to track changes in key gender gaps from baseline to final evaluation. BHA expanded gender requirements, adopting a set of gender indicators. BHA strongly recommends that partners disaggregate all household-level indicators by gendered household type: Adult Female and Adult Male (F&M); Adult Female No Adult Male (FNM); Adult Male No Adult Female (MNF); and Child No Adult (CNA). Additional household or family classification types should also be used, as appropriate and relevant, given the operational, security, and cultural context (e.g., polygamous status, et cetera). Note that applicants are not required to provide sex, age, household type or other disaggregate targets at the application stage; post-award, partners are required to provide disaggregate targets.

Mission indicators: Post-award, the BHA Officer at the relevant USAID Mission will inform the partner about required Mission indicators, which are defined or selected by the Mission. The Mission should provide the partner with the PIRSs that define the indicators. BHA does not expect applicants to include Mission indicators at the application stage.

Context indicators: There are factors in the activity context that are not expected to be influenced by the activity but may affect activity Outcomes. These factors may be identified as LogFrame assumptions. For example, an activity may not implement any intervention to help reduce intercommunity conflict. However, peace and stability in the activity area are necessary to achieve activity Outcomes. For this reason, an activity may want to add contextual indicators to monitor

conflict-related migration, such as "the number of days without access to activity area", or other conflict-related indicators, to have information to assess how much the context may have affected activity Outcomes. Custom contextual indicators that are important to the interpretation of other indicators may also be included. Context monitoring can be done through reviewing secondary data and reports, using qualitative methods such as interviews and discussions, and/or surveys.

Custom indicators: Partners are encouraged to create custom indicators to measure specific, essential activity Outputs, Outcomes, and context for which there are no corresponding BHA or Mission indicators. BHA indicators are developed to meet BHA and USAID reporting requirements. Custom indicators are important and should be carefully identified to enable a partner and the BHA M&E Advisor to track progress along the TOC and to identify how far along a pathway change has occurred. In particular, BHA recognizes that required indicators are not adequate to measure some activity Outputs and Outcomes related to community participation; community assets; resilience; social capital; social accountability; self-efficacy; inter- and intra-community conflict; and governance.

Useful indicators that may be adopted or adapted for use as custom indicators might be found among other U.S. Government (USG) standard indicators (e.g., the USAID Civil Society Organization Sustainability Index) or indicators defined by others (e.g., the United Nations or other donors or professional organizations). If there is a need to develop a custom indicator to replace a similar BHA indicator, discuss and present custom indicators to the AOR and M&E Advisor.

Resources

- BHA's Food Assistance Implementation and Reporting webpage provides links to a variety of resources to assist with the implementation and reporting on food assistance programs.
- The USAID Program Cycle USAID Learning Guide provides basic information and resources related to LogFrames.
- USAID provides guidance in ADS 205 on integrating gender equality and female empowerment in USAID program cycle and, in ADS 204, on integrating environmental safeguards into programs to optimize socioeconomic development results.
- The International Labour Organization's Integrating Gender Equality in Monitoring and Evaluation of Projects discusses the importance of systematically integrating gender equality and a human rights perspective into M&E processes.
- The World Bank's Gender Issues in Monitoring and Evaluation provides ideas for improving the M&E of outcomes and impacts.
- USAID Ending Child Marriage & Meeting the Needs of Married Children: The USAID Vision for Action outlines USAID's efforts to prevent child marriage and to respond to the need of the more than 50 million girls and boys who are already married and have limited access to education, reproductive and other health services, and economic opportunities.
- The United States Strategy to Prevent and Respond to Gender-Based Violence Globally establishes a government-wide approach to addressing gender-based violence and a set of concrete goals and actions for Federal agencies.
- The Food Security Information Network (FSIN) technical paper Qualitative Data and Subjective Indicators for Resilience Measurement discusses basic features of qualitative and perception-based measures for resilience, as well as challenges to using this type of information.

2.3 INDICATOR PERFORMANCE TRACKING TABLE (IPTT)

The IPTT is useful for following activity performance and comparing it against planned progress. Partners should use the IPTT template on the BHA Food Assistance Implementation and Reporting Website to fill out the IPTT with the first M&E Plan submission post-award. (BHA does not require an IPTT with

the application.) The IPTT must be updated and submitted twice annually: (1) with the Annual Results Report (ARR) and (2) with the annual M&E Plan submission two months prior to the PREP.

All quantitative indicators from the LogFrame must be included in the IPTT. Qualitative monitoring inquiries (QM and QS) identified in the LogFrame should not be included in the IPTT. (Results from qualitative monitoring and qualitative studies should be reported on in the ARR, quarterly reports, and any other appropriate deliverables).

The IPTT (and quarterly reports) should also include indicators to monitor environmental impacts. All resilience activities are required to track indicators based on the findings from the Initial Environmental Examination (IEE) and Environmental Impact Assessment (EIA). Partners must develop an Environmental Mitigation and Monitoring Plan (EMMP) that includes the indicators identified to monitor interventions that potentially mitigate negative environmental impact. The M&E Plan must show how the activity will implement the activity's EMMP. There are many performance indicators in the IPTT that also measure activity's interventions that mitigate negative environmental impacts. These are generally agriculture, natural resource management, water, sanitation, and hygiene behavior indicators. Identifying these indicators allows the activity's M&E team to minimize additional data collection burden to monitor indicators identified in EMMP. The EMMP may include indicators that are identified to monitor specific aspects of the environment and cannot be replaced by any performance indicators from the IPTT. These specific custom indicators from the EMMP should be brought into the IPTT, as appropriate.

The IPTT template is an Excel sheet that includes tabs (worksheets) for baseline (BL) indicator values and endline targets; annually reported monitoring indicator targets; base value data sources; deviation narratives/comments; target change comments; and archived monitoring indicators. Partners are encouraged to use the BHA IPTT template, but may use other formats as long as the required information is included. In the baseline/endline indicator and monitoring tabs, the indicators should appear in the same order that they appear in the LogFrame.

Baseline/Endline and Monitoring Indicator Tabs

The purpose of the baseline/endline indicator tab is to identify the values of the BL indicators at baseline (i.e., "baseline value") and the final target values (i.e. that will be measured during the interim/final evaluation). The purpose of the Monitoring Indicator tab is to track values of the annually reported indicators. The following information should be included for each indicator:

Indicator Number: For ease of reference, all indicators in the IPTT must be numbered in the order in which they appear in the IPTT. Mark the first monitoring indicator as 1, the second as 2, and so on, until all indicators have been assigned a number. When an indicator is deleted from an IPTT, its number may not be reused, and the numbers for the indicators that followed it in the IPTT do not change. If an indicator is added to an IPTT, it should be assigned the next number in sequence following the number assigned to the last indicator in the IPTT. For example, if the IPTT previously had indicators numbered 1 through 65, the new indicator would be assigned the number 66. However, its position on the IPTT must correspond to the position of the associated Purpose, Sub-Purpose, Intermediate Outcome, or Output in the LogFrame, i.e., its assigned number will be out of sequence in the IPTT.

Data Collection Method: Enter PBS, RM, or PaBS, as defined above. (Note that Qualitative Monitoring and Qualitative Studies should not be included in the IPTT.)

Indicator: Enter the indicator title, which should match the corresponding Results Statement on the LogFrame and the notation on the TOC. For BHA and Mission indicators, indicator titles must be entered exactly as they appear on the BHA and/or Mission list. No substitutions or modifications will be accepted.

Indicator Type: For all indicators, specify whether the indicator is a BHA (i.e., "BHA BL#" for baseline or "BHA M#" for monitoring), Mission (Mi), custom (C), or EMMP (E) indicator.

- Further identify BHA indicators by number using the BHA Indicator List, e.g., "BHA M4."
- For Mission and custom indicators that have been adapted or adopted from the USG standard indicator list, or a previously archived BHA or BHA indicator list, note the identifier from that list, e.g., "C # (F 4.8.2-26)."
- If an indicator is required by both BHA and the Mission, indicate both BHA (with number) and Mi (with identifier, if relevant).
- If an indicator is part of the EMMP, also indicate E.
- Indicators required by Feed the Future for an activity that are not BHA indicators should be identified as C, with the USG standard indicator number, as applicable. Feed the Future indicators that are also BHA indicators should be specified only as BHA.

Data Source: Briefly describe how data will be collected for each monitoring indicator, for example, maternal and child health and nutrition (MCHN) participants' health cards, savings and loan groups records, health facility records, monitoring form or checklist, organizational capacity assessment tool, household questionnaire, or other appropriate tool for routine or survey data collection. A more detailed description of the data source must be included in the indicator PIRS.

Desired Direction of Change (+/−): For all indicators other than context indicators, enter "+" or "−" to indicate whether the desired direction of change for the value is positive or negative. For example, the desired direction of change for diarrhea prevalence is negative (i.e., a successful activity will reduce diarrhea prevalence), but the desired direction of change in the percentage of farmers who apply an improved technique is positive (a successful activity will promote adoption of the technique among a larger proportion of farmers).

Base Value: Base value must be established for monitoring indicators that refer to activity participants. All output-level indicators related to the activity (e.g., number of individuals trained) typically have "0" as a base value at the start of the award. Other base values such as crop yields should be determined during Y1 using primary data, secondary data, or other existing data sources.

Note: Baseline values refer to baseline indicators and are generated by the research/evaluation partner that is conducting the baseline study and interim/final evaluation; baseline values are population-based for pre-post evaluation designs but may not necessarily be population-based for impact evaluations that use experimental or quasi-experimental

There are many ways base values for monitoring indicators can be established. Partners can use non-probability sampling techniques and use qualitative methods such as group discussions and interviews to collect data to establish base values when no other reliable data are available. In such cases, partners should strategically select communities to adequately represent the diversity of the activity areas. In other instances, partners may be able to collect data to establish base values at the time of registration/census.

The initial IPTT should include a base value for every indicator and required disaggregate. By the end of the refinement phase, the revised IPTT must be submitted with the base value for every monitoring indicator (including Outcome indicators) and required disaggregates.

Base Value Data Source: For all monitoring indicators with a non-zero base value, enter a reference number to direct the reader to text in the base value data source tab that describes the basis of data

collection and the activity-specific groups that will comprise the sampling frame for the base value estimation.

Baseline Value: are the initial value of a baseline (BL) indicator collected as part of the baseline survey (e.g., 34.5% prevalence of stunting among children under five).

Target: The initial IPTT should include targets for every FY and LOA for every indicator, and a interim/final evaluation target for every BL indicator either as a fixed value or an expression in relation to the baseline value. By the end of the second FY (end of the first FY for first-quarter awards), an IPTT must be submitted that also includes a base value for every monitoring indicator, other than contextual indicators, and required disaggregates.

Target Change Comment: Whenever an IPTT is submitted with targets that differ from those in a previous submission, a reason must be provided for every change. An annual target may not be changed in the same or any earlier FY that the actual will be reported, except with the ARR submission for the previous FY. (For example, the last opportunity to change a target for 2021 will be with the submission of the ARR results for 2020.) Changes from "TBD" ("to be determined") to an actual value do not require target change comments. When a reason for a change requires a lengthy explanation, the partner may enter numbered comments on a separate sheet in the IPTT workbook labeled "Target Change Comments" and enter the relevant comment number in the Target Change Comment column on the BL or monitoring indicator worksheet. With every new submission, all comments should be retained on the sheet and numbering should continue throughout the LOA, not start from "1" with every new submission.

Actual Value: With each ARR submission, values measured during the reporting year must be added, as relevant, i.e., every reporting year for monitoring indicators and after the interim/final evaluation for BL indicators.

Percent Difference Achieved for Monitoring Indicators: Divide the difference of the reporting year actual value and the target value, by the reporting year target value and multiply the quotient by 100. Please ignore the direction of change and use the absolute value.

$$\%\ difference\ achieved = \frac{reporting\ year\ actual\ value - target\ value}{reporting\ year\ target\ value} \times 100$$

Table 3: % Difference Achieved Examples

Indicator	Target	Actual	% Difference Achieved	Deviation Narrative/Comment
M2 Number of children under five (0-59 months) reached with nutrition-specific interventions through USG-supported programs	1000	1200	20%	Over 10% difference so required
M4 Percent of households with soap and water at a handwashing station on premises	30	23	23.3%	Over 10% difference so required

For BL indicators, "% *achieved*" calculation guidance: 1) for indicators expressed as percent is the percentage point change; and 2) for indicators expressed as mean/score "% *achieved*" is the percent change from target. More guidance will be provided after the activity receives its endline estimate table from the M&E Advisor.

Deviation Narrative/Comment: When the actual results are 10 percent (not percentage point) above or below the target value, the partner must provide a deviation narrative.

Archived Monitoring Indicators

If/when a partner receives AOR approval to remove an indicator from the M&E system, the partner should: (1) keep the existing data in the IPTT "Annual Monitoring" tab of the IPTT but "hide" the rows; and (2) add the archived indicator to the "Archived Indicator" tab of the IPTT. The data that was already collected for the indicator (prior to it being archived) should be included in the "Annual Monitoring" tab even after the indicator is archived so BHA and the partner may refer back to those data at a later time without having to refer to an outdated IPTT.

The "Archived Indicator" tab is used to record information about indicators that were once on the IPTT but that have been subsequently removed from active monitoring with the approval of the AOR. The template provides columns to record indicator number, source, indicator, and indicator type, as described above. The following should be included:

- **TOC/LogFrame Reference**: Specify whether the indicator was associated with a Purpose, Sub-Purpose, Intermediate Outcome, Output, or assumption.
- **Indicator Level:** Specify whether the indicator measures Output, Outcome, impact, or context.
- **Justification Summary:** Briefly explain why the indicator was deleted.
- **Date Initiated:** The FY for which the indicator was first reported on the IPTT.
- **Date Archived:** The FY for which the AOR approved the deletion of the indicator from the IPTT.

Baseline Values, Base Values, and Targets

In the IPTT, baseline values and base values are **not** the same and these terms should not be used interchangeably because they have different sample frames: **baseline values** refer to baseline indicators and are generated by the research/evaluation partner that is conducting the baseline study and interim/final evaluation; baseline values are population-based for pre-post evaluation designs but may not necessarily be population-based for impact evaluations that use experimental or quasi-experimental methods. By contrast, **base values** only refer to monitoring indicators that are collected and used by the implementing partner.

Baseline Values and Final Evaluation Targets: Baseline values and endline values are entered only for BL indicators, and their values are derived from data collected by the external contractor(s) during the baseline and endline surveys. They reflect the situation at a population level, including households and individuals that are not specifically targeted and do not directly benefit from activity interventions.

Before the baseline survey results are available, partners may enter "TBD" as baseline values on IPTTs and express final evaluation targets in relation to the baseline value, like the targets submitted with the initial LogFrame. Once the survey results are available, partners must include the actual baseline values, as measured by the survey, and replace final evaluation targets with numeric values in all IPTT submissions.

Partners should determine endline targets using an objective approach, which involves determining the following information for each indicator:

- The initial population size at baseline;
- The initial prevalence or count of the indicator (e.g., prevalence of breastfeeding) at baseline;
- The anticipated number of participants or persons targeted for that indicator (e.g., number of mothers / prospective mothers targeted through the intervention);
- The anticipated (target) prevalence or count of the indicator among the cohort of participants identified above at endline;
- The anticipated prevalence or count of the indicator among the rest of the population (i.e., non-participants) at endline;
- The anticipated number of non-participants who may benefit from secondary adoption or indirect exposure to direct participants;
- The anticipated overall prevalence or count of the indicator at the population-level (i.e., including participants and non-participants); and
- Population growth or other demographic trends that may contribute positively or negatively to endline, population-level changes in prevalence rates or totals.

The BHA Food Assistance Implementation and Reporting webpage includes a toolkit with an objective target-setting tool that implementing partners may use to help set targets for baseline/endline indicators using this approach.

BHA recommended interim/final evaluation target for BHA "BL" indicators:

- An average minimum of 2 to 2.5 percentage point reduction of prevalence of stunting per year of implementation
- An average minimum of 3 to 4 percentage point reduction of prevalence of underweight per year of implementation
- An average minimum of 4 percentage point reduction of depth-of-poverty per year of implementation

Base Values and Annual and LOA Targets: Base values and LOA targets are measured by the partner for monitoring indicators. They count Outputs and measure Outcomes among those who directly benefit from related activity interventions. For every monitoring indicator, the IPTT must include a base value and targets for every FY and the LOA.

For most Output and some Outcome indicators, base values will be zero, however some base values are non-zero (e.g. agricultural yields). Another example of a non-zero base value may be when a partner has been working in the same area prior to a new award and supported some community groups and associations with funds from a previous USG award. If the partner plans to continue to support these existing groups under the new activity, in addition to other groups, then the base value for the Output "Number of institutional settings gaining access to basic drinking water services due to USG assistance," for example, would be a count of the previously supported institutions that will receive continued support. For these indicators in the first and all other IPTT submissions, zeros or a count of "continuing" entities at activity start-up should be entered as base values, and numeric annual and LOA targets must be established.

Base values and LOA targets for annual Outcome monitoring indicators should reflect the desired and actual values measured only among the relevant participant groups, not the situation among all activity participants. For example, measures of child feeding practices should consider only those who benefit directly from interventions that pass messages about or provide support to enable optimal feeding practices, and measures of incremental agricultural sales should be collected only from smallholder,

direct participants of activity interventions intending to increase sales of the targeted value chain commodities.

For many annual Outcome indicators, a partner will need time to collect data to establish base values that truly represent the situation among activity participants as interventions begin. These include indicators related to such things as prevalence of attitudes or practices, agricultural production or sales, and anthropometric measures. To allow time for the enrollment of participants and for the collection of information from them, in the early submission of the activity IPTT, "TBD" may be entered for these indicators' base values, LOA targets, and all annual targets in between. Collection of data to establish true participant estimates should be completed within approximately a year, and partners must submit an IPTT with real estimates and targets with the ARR of the first FY if the award was awarded during the first quarter of that year and with the ARR of the second FY for those awarded after the first quarter of the first FY.

The magnitude of change among participants during the LOA is expected to be greater than the magnitude of change in the population. Therefore, LOA targets for monitoring indicators should always be more ambitious than final evaluation targets. In addition, to demonstrate a population level change, LOA targets for monitoring indicators should always be much higher considering the coverage of an activity in the target communities. The smaller the coverage within each community, the greater the LOA target should be otherwise the activity will not be able to demonstrate a population level change.

The sources of the base value should be indicated in the "Base Value Data Source" in the IPTT, with additional detail in the "Description" section of the indicator's PIRS.

Disaggregates

Partners must include baseline, base, and target values for all required disaggregates of BHA indicators that are defined in the PIRS. This includes:

- Baseline values and final evaluation targets for all required disaggregations of BHA BL indicators;
- Base values and targets for all required disaggregations of BHA monitoring indicators for the LOA and all years other than the FY of the award; and
- Targets for all required disaggregations of BHA monitoring indicators for the first FY of the award for activities awarded during the first quarter of the FY.

New and Continuing: For many interventions, activities include some new participants on an annual basis while other participants who joined the activity in a previous year continue to participate through the reporting year. For example, a partner may train 1,000 farmers in Y2 and then re-train those same farmers in Y3 while also training an additional 500 farmers. In Y2, there would be 1,000 "new" farmers trained; in Y3, there would be 500 "new" farmers trained and 1,000 "continuing" farmers trained. For other indicators, such as the number of people gaining access to safe drinking water, the activity should monitor access on an annual basis to confirm the number of continuing individuals rather than generate an estimate based on an assumption that all water points installed the previous year are functional. Only unique participants who joined the activity during the reporting year should be reported as "new". Therefore, the LOA target should be the sum of all new participants. A participant should not be counted multiple times to compute LOA target and actuals. The indicator PIRS identifies whether the indicator must be disaggregated by New and Continuing. (Note that BHA previously classified indicators as either "cumulative" or "noncumulative." The "new" and "continuing" disaggregation has been adopted to simplify indicator tracking and reporting.)

IPTT Submission and Revision

The deadline for submission of the first IPTT will be specified in the award document or communicated to the partner by the AOR. The IPTT will be updated and submitted at least twice per year: (1) with every ARR submission; and (2) with the M&E Plan submission two months prior to the PREP submission.

With the ARR submission, the partner must provide actual values for the reporting (i.e., just-completed) FY. With each ARR, the partner may request changes to targets for the current year (i.e., the year following the reporting year). However, once a current-year target has been approved in that year, it may not be changed. This is the value against which actual achievement will be compared in the next ARR. If circumstances change that justify under- or overachievement of this target, the partner will have opportunity to explain what happened in a Deviation Narrative/Comment in the IPTT submitted with the next ARR.

At any time, with justification, the partner may request approval for other modifications to the IPTT, e.g., the addition or deletion of indicators; data collection methods; or targets for future years, LOA, or final evaluation.

All requests for revisions to the IPTT must include a narrative that describes and justifies the proposed changes. The addition, removal, redefinition, or change of data collection methods for an indicator on the IPTT often requires changes to other components of the M&E Plan.

2.4 DATA COLLECTION SHEETS

Data collection sheets refer to performance indicator reference sheets (PIRS) for quantitative indicators and qualitative inquiry planning sheets (QuIPS) for qualitative inquiries. The sections below provide information on what should be included in the PIRS and QuIPS. Note that it is mandatory to provide a PIRS for all quantitative indicators included in the LogFrame or IPTT; and it is mandatory to provide a QuIPS for all planned qualitative monitoring efforts and stand-alone qualitative studies during the activity implementation stage (i.e., a QuIPS should not be used for formative research studies conducted during the first phase of R&I activities).

2.4.1 PERFORMANCE INDICATOR REFERENCE SHEETS (PIRS)

A PIRS should unambiguously define a quantitative indicator and describe how raw data will be collected and used to calculate the indicator, the methods and frequency of collection, and the calculations used to derive final values and disaggregates. The PIRS must describe the indicator to such a level of detail that anyone could use the PIRS to know exactly:

- What raw data are needed;
- What questions to ask or processes of observation to follow to get the accurate raw data without causing harm;
- Who will collect the data or observe the activity intervention;
- Which tools will be used to collect the data;
- From whom data will be gathered, what performance will be assessed, and/or what infrastructure or activity intervention will be observed;
- Precisely when the data will be collected;
- How the collected data will be used to calculate the indicator value; and
- In what unit the data will be presented and the range of possible values.

Draft PIRSs for all custom baseline and interim/final evaluation indicators must be submitted to the AOR no later than two (2) weeks prior to the baseline workshop. The final PIRSs for these indicators, revised on input received during the baseline workshop, must be submitted within 14 days of the end of the workshop.

PIRS should be submitted with the first M&E Plan submission and with every subsequent M&E Plan submission (two (2) months prior to PREP submission). If additional updates need to be made outside of these annual submissions, the partner may submit them to BHA at any time. The partner should contextualize PIRSs for all BHA and Mission indicators to the activity context and activity-specific information. The AOR will give final approval of the partner's M&E Plan only when it includes a PIRS for every indicator in the IPTT.

With every request for a revision to the IPTT that adds or modifies a monitoring indicator, the partner must also submit to the AOR a modified PIRS and a narrative justification for the requested changes.

BHA and Mission Indicator Sheets

Each BHA and Mission PIRS standardizes the meaning and derivation of an indicator within and across activities. Partners must collect data for BHA and Mission indicators using the definitions, questions, and instructions provided in the PIRS. The BHA Indicators Handbook provides PIRS's for most BHA indicators:

- The BHA Indicators BHA Indicator Handbook Part I: Indicators for Baseline and Endline Surveys for Resilience Food Security Activities contains PIRSs for BHA indicators required for collection in baseline and endline surveys.
- The PIRSs for BHA monitoring indicators are in the BHA Indicator Handbook Part II: Monitoring Indicators.
- Performance Indicator Reference Sheets for Mission indicators should be obtained from the BHA Officer at the Mission post-award.

For some BHA or Mission indicators, partners must add activity-specific information to the BHA or Mission PIRS to clarify details about the indicator. For example, for the annual indicator, "Number of individuals in the agriculture system who have applied improved management practices or technologies with USG assistance," the activity's PIRS should add text after each technology type to specify the different technologies and practices that will be promoted and counted, e.g., which type of crops and seeds and which specific crop and livestock management practices. Additional text should also clarify the types of "individuals in the agriculture system" who will be participants and counted for each technology type, e.g., characteristics of individuals who will be targeted, types of entrepreneurial processors, and traders of which commodities. It is helpful to identify the interventions that will benefit these individuals (i.e., the intervention groups).

For all BHA and Mission monitoring indicators, specific activity intervention groups or Outputs that will comprise the sampling frame for the indicator must be identified as such. For example, for the indicator, "Number of people trained in disaster preparedness with USG assistance," in the "Definition" box, the partner should insert text that identifies the specific interventions in which the participants participate and when they will be counted, something like: "Participants who will be counted include (1) Community Resilience Committee members who completed a series of three training sessions related to resolution of inter-community conflicts and/or prepared the community to recognize and respond to pending drought, and (2) community advocates who completed the activity training module designed to build capacities to organize and moderate discussions between members of marginalized groups and the larger community." This example shows the specific training module that an individual must complete to be counted for this indicator and the types of trainees. In this case, the types of trainees are defined by

their role in the community. For BHA indicators, the text clarifying the sampling frame would fit well in the "Measurement Notes: From Whom."

Partners may make other additions to clarify the use of a BHA or Mission indicator in the activity's M&E Plan. For example, text may be added to identify the indicator as part of the activity's EMMP and explain how the indicator is environmentally sensitive to the activity context. Clarifications inserted in the PIRSs do not "change" the BHA or Mission indicator; they simply add more information. **All additional or modified text in an BHA or Mission PIRS should be highlighted to clearly differentiate it from the standard PIRS text.**

The partner may not change the core definition or meaning of a BHA or Mission indicator. For example, if the indicator counts individuals, the partner may not count groups. If the BHA indicator specifies a count of infrastructure that was completed during the year, the partner may not change to count infrastructure on which any work was done, whether completed or not. If a partner wishes to substantially change the core definition or meaning of a BHA or Mission indicator, the partner should treat it as a custom indicator (and label it as such).

A common way that partners have transformed a BHA indicator into a custom indicator is by using a BHA BL indicator as a monitoring indicator. Partners collect raw data using the same questions as those used in the baseline, and they perform the same calculations defined in the PIRS. However, because monitoring indicators measure results only among direct participants, the sampling frame for the custom monitoring indicator (the specified participant groups) is different from the sampling frame for the BHA BL indicator (population). Therefore, the indicator has "changed" and is no longer a BHA indicator.

Custom Indicators

Partners must develop and submit a PIRS for every custom indicator following the BHA template provided in the *BHA Indicator Handbook II*. An indicator's PIRS should fully describe the meaning of the indicator value in unambiguous terms, and it must include details about exactly what raw data must be collected to calculate the indicator's value; when, where, by what mechanism, from whom, and by whom the raw data will be collected; and how the raw data will be aggregated and used in calculations to derive the final indicator value. These details should be sufficient so that the reader could derive the indicator's value using the information in the PIRS. The different sections required for a custom PIRS are described below.

Indicator: Indicator number and name. The unique number associated with the indicator in the IPTT PIRS should match the indicator numbers included in the Logframe.

Definition: A precise description of what the indicator is and how it is derived. This section should unambiguously define key words, terms, and phrases.

As an example, for an indicator "% of children who completed postnatal visits on time," the key words that require definition are "children," "completed," "postnatal visit," and "on time." Thus, it would identify criteria for a "postnatal visit," e.g., it might specify who is qualified to provide postnatal care (e.g., nurse, doctor, or midwife) and what care must be given (e.g., immunizations, measurements of length and weight). It must also explain how many visits should take place and/or what services must be received for "completion." The definition of "on time" would indicate the maximum amount of time that may lapse after the child's birth before the visits are "completed." "Children" should be defined as those who benefit directly from interventions that promote postnatal visits and who are old enough to have completed all the postnatal visits, but who are not so old that they were long past the time when the visits should be completed. Therefore, if postnatal care should be completed within six (6) months of delivery, "children" could be defined as those whose mother is a member of a care group or recipient of

supplementary feeding and who reached the age of 6.5 months (195 days) during the reporting period (allowing an additional two (2) weeks for completion).

For counts of Outputs, individuals, or other entities that meet a certain set of criteria, the criteria for being counted should be clearly laid out in the definition. For example, for an indicator, "Number of communities with multi-sectoral development plans proceeding on schedule," the definition should identify the number and types of sectors that must be included for the plan to be considered "multi-sectoral" and the basic requirements of the plan (e.g., written document, endorsed by the community and regional authorities) and lay out criteria for being "on schedule" (e.g., completed x% of annual action items between October 1 of the previous year and September 30 of the reporting year).

For data collected using questionnaires, the indicator definition should include the specific question(s) that data collectors will use to gather the raw data needed to calculate the indicator and its disaggregates. For the first example above, to select the right children and get an accurate count, the necessary raw data might be the mothers' participant registration number, the children's age, the dates of every postnatal visit, and the types of services received at each visit. Note the need to define not only data used to calculate the indicator, but also data to ensure selection of participants who are in the defined sampling frame (e.g., age).

Note that the definition for custom indicators should **not** explain why the indicator is included in the IPTT, i.e., do not provide a rationale for either the indicator or the interventions for which it measures performance.

For indicators that require a calculation, show the formula for calculation and fully define each factor in the formula, e.g., identify the characteristics of individuals who will be counted in the numerator and in the denominator for all proportions or explain how a score sheet will be converted to an aggregate rating score for a health center's performance to be rated as "outstanding," "acceptable," "partially acceptable," or "unacceptable." For the example related to postnatal visits presented above, this text could define the numerator as "the number of children who completed postnatal visits on time" and the denominator as "the number of children." Note that because "children," "postnatal visits," "completed," and "on time" are defined precisely in the Definition section, there is no need to repeat the participant status, age limitation, timing, et cetera.

How to Count LOA: For all indicators, clearly describe LOA value or how it will be calculated. Specify if the LOA is a unique number.

Unit The unit of measure in which the indicator will be presented (e.g., number, percent, kilometers, acres, or score) and, as applicable, the range of minimum and/or maximum indicator values or a list of the possible encoded values, with their meanings. For indicators reported as a "score," for example, a child dietary diversity score, an explanation of the method of scoring must be included in the Definition section.

Disaggregate by: A list of all the different ways the indicator values will be disaggregated (e.g., male/female, gendered household type, polygamous household type status, youth/adult, urban/rural, region, ethnicity, region of origin, age group, type of training, religion). All USAID activities must collect appropriate sex-disaggregated data for all people level indicators, ask clear questions about gender roles to reveal both intended and unintended positive or negative changes, and develop indicators designed to track changes in key gender gaps from baseline to final evaluation. BHA expanded gender requirements, adopting a set of gender indicators. BHA strongly recommends that partners disaggregate all household-level indicators by composition of household: Adult Female and Adult Male (F&M); Adult Female No Adult Male (FNM); Adult Male No Adult Female (MNF); and Child No Adult (CNA). Additional household or family classification types may also be used, as appropriate and relevant, given the

operational, security, and cultural context (e.g., polygamous status, et cetera). Be sure to define any key terms used in the disaggregation section, if necessary.

Note on disaggregation of percentages: The categories of disaggregation should define the numerator and denominator. For example, disaggregation of a percentage by sex would calculate the percentage as follows:

- Female: $\frac{\#\ of\ female\ farmers\ who\ applied\ a\ technique}{\#\ of\ female\ farmers}$
- Male: $\frac{\#\ of\ male\ farmers\ who\ applied\ a\ technique}{\#\ of\ male\ farmers}$

Clearly define what is new and continuing for each New/Continuing disaggregation within the indicator context.

Level: Identify the indicator as an Impact, Outcome, Output, or Context indicator. Note that the level must correspond to the related level in the TOC, i.e., Outputs must be associated with Outputs, and Outcomes must be associated with Purposes, Intermediate Outcomes, or Goals. Impact indicators typically will relate only to Goals. Context indicators could appear at any level or be associated with an assumption.

Direction of Change: For output, outcome, and impact indicators, indicate the desired direction of change using plus or minus signs (i.e., "+" or "-") or plain language, e.g., "higher is better."

Data Source: Describe how data will be collected for each indicator, for example, maternal and child health and nutrition (MCHN) participants' health cards, savings and loan groups records, health facility records, monitoring form or checklist, household questionnaire, or other appropriate tool for routine or survey data collection. Be as specific as possible and avoid generic terms like "project records."

For BL indicators, this will always be "Baseline and Interim/Endline surveys."

Note that extraction of information from an internal report is **not** a method of collecting data. Data in the report were collected by activity staff or participants before the report was written, and extraction from the report is a later step in the data flow. An exception would be information for context indicators drawn from a report generated **outside** of the activity. For example, information could be taken from a market report compiled by the government using data collected by others. In such cases, for the purposes of the activity, "Extraction from the Ministry of Trade quarterly market report" would be the data collection method, because neither activity staff nor participants were involved in the collection of the data for the report.

Foreign Assistance Standardized Program Structure (SPS): Include SPS indicator reference number, if applicable.

Measurement Notes: Expand on the basic definition provided in the IPTT about the sampling frame, the frequency of data collection, and the method. The sections on BL indicators will be the same, with fixed information entered as described below. For monitoring indicators, the information will vary in each of the subsections.

Who Collects: Identify who will collect data (e.g., the implementing partner or an external, third party data collection firm. For BL indicators, enter "Research/evaluation partner." For monitoring indicators, describe, by function or role, the individual(s) who will record the raw data at the site of collection. Examples for annual indicators include M&E field monitors, community health volunteers, activity health

and nutrition field supervisors, food distribution receptionists, vendors who redeem vouchers, and agriculture extension workers.

From Whom: Describe which individuals, households, or communities will be counted as part of this indicator and/or from whom data will be collected. This is typically either activity participants or the population in the BHA activity implementation area. For BL indicators, it is "Population." For monitoring indicators, identify the specific category(ies) of participants (i.e., intervention groups) or Outputs that will be counted, questioned, or observed. Examples include food for asset infrastructure Outputs completed during the year, training sessions completed during the year, FFS participants, women's group members, or village resilience committee members.

Method: This should include information on data collection approach, sampling design and sampling frame and the frequency of data collection to be used (e.g., population-based survey, participant-based survey, routine monitoring). For BL indicators, this will always be "Population-based Survey" for performance evaluation. For impact evaluation, a "Participant-based Survey" could also be possible. For monitoring indicators, it will either be Routine Monitoring or Participant-based Survey.

Frequency of Collection and Reporting: Describe when data will be collected, e.g., quarterly, ongoing, or within 30 days of a distribution; and when data will be reported (usually annually). For BL indicators, the frequency is "Baseline and interim/final evaluation." For monitoring indicators, identify how often the raw data collection takes place. Examples include "Monthly when savings and loan groups meet," "Annually, at the end of rice planting season," "Semi-annually in January and July," "After the first and second harvest," and "Mid-quarter at FFS sessions." Frequency of collection is not necessarily the same as reporting frequency. For example, while training data is collected monthly, it is reported quarterly or annually.

Base Value Info: Describe how base values will be determined. Note that all output indicators should have base values of zero; some outcome indicators may have non-zero base values.

Reporting Notes: This section provides guidance on how to report the indicator in the IPTT. The section includes specific instructions on reporting, especially important for indicators with multiple levels of disaggregates.

Further Guidance: References to additional information or guidance for an indicator that has been defined elsewhere. For example, if the activity is adopting an internationally recognized indicator that has not been adopted by BHA, provide links to associated reference documents or publications here.

Note that BHA no longer requires the PIRS to indicate if an indicator is cumulative or non-cumulative.

2.4.2 QUALITATIVE INQUIRY PLANNING SHEETS (QUIPS)

A Qualitative Inquiry Planning Sheet (QuIPS) is used to outline the methods and plans for conducting qualitative monitoring (QM) and stand-alone qualitative studies (QS) that are part of an activity's performance monitoring system. (Note that the QuIPS should only be used for qualitative inquiries that take place during the activity implementation phase. By contrast, scopes of work should be used to plan and communicate to BHA the methods that will be used to conduct qualitative, formative research during the refinement phase of the activity.)

Annex V includes a template for the QuIPS. The components of a QuIPS are as follows:

Title and Type (QM or QS): Insert the title of the qualitative study or monitoring inquiry. Indicate whether the inquiry is a qualitative study (QS) or qualitative monitoring (QM) and provide a number that corresponds to the order in which it appears in the LogFrame. For example, the first qualitative monitoring inquiry in the LogFrame should be labeled "QM 1" in the LogFrame and in the QuIPS.

Purpose and Objective: In a few lines, describe how the information being collected will inform programming or decision-making. This should include a statement that summarizes the purpose of the study, using the following illustrative format: The purpose of this qualitative (monitoring or study) is to (understand, explore, describe) the (phenomenon of interest) of/for/among (activity participants, population of interest) at/in (study site). Examples include:

- The routine qualitative monitoring inquiry will annually assess the attitudes and changing practices around climate change adaptation strategies among communities participating in the activity. Anecdotal evidence provided by staff suggests that while farmers generally recognize that the weather has been different from when they were children, there is not a widespread acceptance that climate change patterns will continue in the future. The purpose of this qualitative monitoring is to understand how to tailor the climate adaptation communications strategy and agricultural trainings.
- The purpose of this qualitative monitoring is to understand why maize crop yields are higher among male farmer participants compared to female farmer participants in spite of roughly equal application of improved farming practices across female and male farmer participants. The first farmer survey conducted in [month, year] showed that maize yields were significantly higher among male farmers compared to female farmers even though both males and females were applying at least 3 key improved practices. This qualitative monitoring is intended to (1) better understand which practices are being consistently applied by female and male farmers; (2) better understand what external factors may be contributing to low yields for women famers, such as more limited access to high quality seeds; and (3) inform the re-design of the agriculture interventions on the activity to improve crop yields for women in particular.

Research/Inquiry question(s): Outline the high-level question(s) that will drive the monitoring or study efforts. (This is not a list of questions to be asked in an interview or focus group, but rather the broader questions that will guide the qualitative inquiry.) Using the examples above, this might include the following:

- Why are diet diversity and feeding practices generally better among children of first wives of polygamous households compared to second or third wives? Why are diet diversity and feeding practices roughly equivalent for first children of polygamous households and all children of all non-polygamous households?
- Why were crop yields lower among female farmers compared to male farmers even though both groups tended to apply at least three improved practices? Which practices are female farmers applying? Which practices are male farmers applying? What other factors may account for low crop yields for female farmers?
- Why do community members generally reject the concept of climate change even though there is widespread acknowledgement that the seasons have been shorter and the rains have been coming later for more than 15 years? How do community members perceive the concept of "climate change"? How might this differ from their perceptions of systematic changes over time in the season length and timing of rainfall?

Additional example research questions include:

- What factors or circumstances hinder or facilitate behavior change around handwashing?
- Why are there substantial differences in uptake of healthy eating between Region X and Region Y?
- To what extent are WASH interventions effectively linked with other activity interventions and/or complementary public and private health, nutrition and social services in the community?
- Why do some households and communities adopt new technologies introduced by activity interventions and effectively manage climate shocks/risks, while others do not?

Question/inquiry type (Process/Output/Outcome/Impact/Context/Cross-cutting theme): Define the type of questions being asked in terms of how the question relates to the activity TOC. For example, qualitative monitoring on women participation and engagement in Care Group lead mother trainings may address multiple levels in the TOC that relate to gender dynamics, child nutrition outcomes, and pre- and post-natal care outcomes. A stand-alone study to explain why adoption of handwashing practices is lower in Region X compared to Region Y may, by contrast, only relate to a single outcome related to handwashing in the TOC.

Qualitative Design and Methodology: This section includes the following:

Data source(s) and research methods: Indicate the source of information (e.g., intervention groups, female participants, youth) and the methods that will be used, such as focus group discussions, key informant interviews, participant observation, or other. If appropriate, indicate if interviews and/or group discussions will be conducted at multiple levels (e.g., community field workers, district or regional stakeholders). Using one of the examples above, this may appear as follows:

- This qualitative study will use focus groups, key informant interviews, and observation to understand differences in adoption of key agricultural practices between female and male farmers.

Sampling strategy: Briefly describe sampling approach for site and respondent selection. Describe how respondents will be identified for inclusion/selection in the study or monitoring practice. Clearly define the characteristics of individual or household level respondents that should be included in the study. If appropriate, define what intervention groups the respondents will be a part of (e.g., mother care groups, farmer field schools). For a study, participants may be identified for participation through a form of purposive sampling (such as quota sampling or snowball sampling). For monitoring, they may be participants of a specific intervention with varied characteristics. Using one of the examples above, this may appear as follows:

- Equal numbers of female and male farmer participants will be invited to participate in focus group discussions, which will be held separately (i.e., female-only focus groups and male-only focus groups). At least two (2) pairs of focus groups will be conducted in each of the five activity intervention areas, which is a total of 20 focus groups. Follow-up, individual key informant interviews will be conducted with Producer Organization (PO) leaders (who are primarily male) as well as female and male PO members, as needed. Using the most recent agricultural survey data, field observations will also be conducted to observe differences in practices such that high-adoption and low-adoption individuals' fields will be observed both during the field preparation period in May and during harvest in September.

Data collection tools, and implementation plan: Specify the type of qualitative tool(s) that will be used (e.g., topical outline, mapping, matrix ranking, Venn diagrams) and describe how the tools will be tailored to different categories of respondents. Briefly describe the specific topics and/or indicators around which information will be collected and analyzed. Ensure the tools align with the selected methods. Indicate how the data will be collected (e.g., by field monitors, external consultants). Using the example above, this may appear as follows:

- Focus group guides will be used to guide the focus groups; a lead facilitator and note-taker will conduct each focus group. Key informant interview guides will be used to guide PO leader and member interviews; a lead interviewer and note-taker will conduct the interviews together. Field staff and a short-term agricultural/gender consultant who will be leading the study will conduct field observations. The short-term agricultural/gender consultant will develop the interview guides and train the facilitators, note-takers, and field staff who will be collecting data.

Both female and male facilitators will be used to ensure women and men participants feel comfortable in the discussions and interviews.

Frequency and timing: Describe the frequency and timing of the data collection, such as routine monitoring (monthly, quarterly, annually) or a discrete study at a crucial point of the activity. Explain how the timing of the study aligns with the award cycle and monitoring system, as well as other factors such as agricultural cycles or other seasonal or special events (e.g., holiday periods, weather patterns, elections). Using the example above, this may appear as follows:

- Focus group discussions and key informant interviews will take place during the crop preparation cycle; as such, these will need to be limited in time due to the time constraints people face during this time of year. Field observations will be conducted during the field preparation in May as well as during harvest in September.

Training requirements: Describe the timing of and topics that will be covered during training for staff or contracted interviewers who will be collecting and/or analyzing data, to ensure that they understand the areas of inquiry, ethical considerations, and are able to adapt and make informed decisions in the field, in consultation with M&E and study team leads. (If this is included in the M&E Plan Staffing and Capacity Development Plan, include a note in the QuIPS to refer to that section of the M&E Plan.) Using the example above, this may appear as follows:

- An external consultant, specifically a short-term agricultural/gender expert, will oversee the study and train the facilitators, note-takers, and field staff who will be collecting data. The Activity M&E Plan provides additional detail on the standard data collection training that all staff receive prior to their participation in data collection activities.

Data recording, data management, and quality assurance: Describe the system for data entry, file naming, review, storage, data security and communication. Define roles and responsibilities of staff, interviewers and data recorders. Describe steps that will be taken to ensure quality data collection and recording, such as: adequate time for training, data collection and data entry; daily debriefs; supervision in the field; data review by M&E staff or study team leads. (If this is included in the M&E Plan Monitoring Section, include a note in the QuIPS to refer to that section of the M&E Plan.) Using the example above, this may appear as follows:

- There will be a dedicated note-taker for all focus group discussions and interviews. The discussions and interviews will not be audio recorded due to cultural taboos and privacy concerns raised during the refinement stage. As such, the field notes from the facilitator and the detailed notes from the note-taker will comprise the primary data that will be used for the group discussions and interviews. Field notes from field observations will also be used. Note-takers will use a unique ID code system to record the location and identify of study participants, such that their identify will be recorded in a separate, secure tracker in CommCare and field notes and observations will only include participant IDs and no other unique identifiers or personally identifiable data. Staff will be trained and supervised by the external, short-term agricultural/gender expert on how to treat the data and use the ID system to ensure privacy of participants.

Data Analysis Plan: Describe how data will be analyzed and/or synthesized with existing monitoring data or secondary sources. Typically one of two methods is used to analyze primary, qualitative data: (1) a Manual matrix approach, which involves using Microsoft Word or Excel to organize and synthesize data according to established themes and categories; or (2) specialized qualitative data analysis software is used to analyze the data. For the scope and depth of qualitative monitoring and qualitative studies for BHA resilience activities, the first approach will likely be more appropriate given that it is quick, user-friendly, and facilitates iterative data analysis.

Disaggregated / Categories: If applicable, identify any categories or disaggregations that may be used in the analysis. For example, if the study question is comparing female and male farmers from different regions, this section might articulate that the analysis will be looking by gender across regions.

Deliverable: Identify where the findings from the qualitative monitoring or study can be found once it has been completed. Using the example above, the deliverable may be a half-page summary of the study in the next ARR. **BHA does not require a stand-alone deliverable for qualitative monitoring or qualitative studies that are conducted as part of the broader monitoring system.** A deliverable can take multiple forms, such as operational notes or guidance, a section of the ARR, a dashboard, presentation, or formal report or publication. If a partner has a question about what is expected as a deliverable for a qualitative inquiry, contact the BHA AOR and M&E Advisor for further guidance.

Utilization/Application: Describe how the findings or information from the qualitative monitoring or study will be used, and by whom. Describe any plans to share findings with participants, communities, and other relevant stakeholders.

Limitations and Risks: Identify any known limitations and risks that may affect the monitoring activity or study, or compromise the quality of the findings. This may include, for example, security constraints; limitations of time, resources, information; researcher bias or capacity; or potential risks to respondents of participation.

Ethical Review Status/Informed Consent (*for Qualitative Studies that include direct data collection from human subjects only): Confirm whether an ethical review was done and any plans or status of ethics approvals (e.g., an institutional review board). If an ethical review was not conducted, explain why it was exempted, for example because the information was being collected from program participants, posed minimal risk, and did not meet the organization's definition of "research". Confirm that an informed consent process will be used for data collection of all kinds (e.g., participant observation, interviews, audio or video recording, photographs). Note that the partner must comply with all informed consent guidance in ADS 579.

Implementation Timeline: Briefly explain the timeline for data collection, data processing, and information sharing and utilization.

Resources

- The BHA Indicators Handbook Part I: Indicators for Baseline and Final Evaluation Surveys contains PIRSs for all current BHA baseline and endline indicators.
- The PIRSs for BHA monitoring indicators are in the BHA Indicators Handbook Part II: Monitoring Indicators.
- The Gender Indicators: What, Why and How? briefly highlights the use of gender indicators
- The Asian Development Bank's Toolkit on Gender Equality Results and Indicators provides information on monitoring and evaluating gender equality results.
- USAID provides guidance on the IEE and EMMP provides environmental safeguards and compliance, including IEE and EMMP development.

CHAPTER 3. MONITORING STRATEGY

3.1 ACTIVITY MONITORING

3.1.1 OVERVIEW

The Monitoring Strategy is a required section of the M&E Plan that should describe in detail the processes, systems, tools, and actors involved in the collection, processing, and use of data. The plan should articulate how the activity will use quantitative and qualitative performance monitoring data to monitor the activity's performance and theory of change; and how it will use secondary data or information to monitor conditions external to the activity such as environmental, security, cultural, or other factors that may affect implementation.

The Monitoring Strategy should provide a comprehensive overview of how indicator data and qualitative information will be gathered, synthesized, and used to monitor performance of the activity and the operating environment. The Monitoring Strategy should align with the data collection methods described in the PIRS and QuIPS, and outline precisely how data and information will ultimately make it from the enumerator/data collector or activity field officers into a summary report, dashboard, or other resource that will be used by internal technical teams to monitor activity performance and context.

A preliminary Monitoring Strategy must be submitted with the application as part of the abridged M&E Plan. Post-award, partners will develop the Monitoring Strategy during the refinement phase and then submit the plan as part of the first full M&E Plan submission.

For partnership and consortium-managed awards, the Monitoring Strategy must clearly describe the roles and responsibilities for data collection, transfer, safeguarding, and use for all consortium members.

The Monitoring Strategy should also describe the analysis plan, data quality assurance approach (including data flow and level of aggregation), utilization, and how the partner will establish base values. This may include how the raw data will be checked for quality, how the data will be analyzed and who would analyze them, how the summary reports then go to a data entry team who uses them to update information in a database containing historic information about the groups, or how an M&E analyst will access the database and qualitative findings to generate values for several IPTT indicators using the calculations in the PIRS.

Partners may use a data flow diagram, matrix, or other form of summary presentation to show, at each step along the way, when, where, how, and by whom raw data are collected, aggregated, processed, and disaggregated to achieve the annual reporting values. This summary may also show how robust qualitative insights fit into the activity's overall monitoring strategy. The methods, timing, and responsible parties for the various steps of data collection may differ for different data.

3.1.2 QUANTITATIVE INDICATOR DATA COLLECTION

Most BHA monitoring indicators measure changes among all or a sample of direct participants[5]. Quantitative data for other custom, annually reported monitoring indicators are sometimes collected at the community or population level. For BHA and Mission indicators, direct participants include individuals, households, communities, institutions, and groups that receive significant goods, services, and/or other support as a direct result of activity interventions. In cases where an intervention employs a deliberate strategy of training a small number of individuals to deliver services or pass knowledge directly to other individuals or organizations (e.g., cascade training), both those who are directly trained by activity personnel and the individuals who are then trained by those (directly trained) individuals or

[5] In some cases, indicators measure outputs such as kilometers of roads built.

organizations, in accordance to an activity-defined training or knowledge transfer plan, are considered direct participants. People who might be exposed to activity messages by attending an occasional sensitization meeting or community intervention such as a theater presentation, hearing a radio message, or viewing a poster in the health clinic or input suppliers shop can be considered "indirect beneficiaries" of the activity and **should *not* be counted** as direct participants for BHA or Mission indicators.

In general, secondary adoption of activity-promoted practices and behaviors after informal contact with an activity participant is expected but since the activity did not invest significant time or other resources on indirect beneficiaries,[6] these individuals should not be sampled for participant-based surveys. Additionally, neighbors and other household members who apply new practices based on observation and/or interactions with direct participants should also be excluded from BHA monitoring survey sampling frames. (There are a few instances when people who directly benefit from an activity Output, such as community infrastructure developed or rehabilitated by the activity, an activity participant who received support from an activity and created employment for others will be counted under employment monitoring indicator are considered direct participants even though they may have had no direct contact with any activity staff or trainees of the activity. For example, those individuals who live near a rehabilitated borehole and are assumed to be drawing water from it are counted for the indicator "Number of people who gained access to an improved drinking water source," regardless of the level of their exposure to other interventions. The relevant indicators' PIRSs make it clear when people with little or no exposure should be counted.)

Ideally, monitoring indicator values should be derived from data collected from a probabilistic sample or from all direct participants who participated in the intervention(s), not a subset. In cases where the indicator's sampling frame is only a subset of direct participants, the sampling methodology and selection process must be approved by BHA. In these cases, the sampling frame and selection process must be described in the Monitoring Strategy and PIRSs.

Routine Monitoring and Surveys

There are two general approaches for collecting quantitative data on BHA indicators: routine monitoring and surveys. The choice of data collection method should be driven by the type of indicator and when, and how, indicator data will be used by activity technical staff to make adaptive management decisions. Collecting data using routine monitoring using program staff is typically faster than a survey, the analysis is relatively simpler, data collection cost is cheaper, and the process empowers program staff. To take advantage of these benefits, BHA strongly encourages partners to use routine data collection when possible.

Routine Monitoring (RM) from all activity participants: This typically involves implementing staff members getting information for all participants of a given type at the intervention sites. Examples of routine data collection are:

- Site visits (e.g., to observe plant spacing and yield).
- Activity records (e.g., participant training registration sheets) or government records (e.g., number of pregnant women visiting health facilities).
- Trained service providers, e.g., extension agents or activity staff, record information about trainees, participants in FFSs, or mothers who attend mothers' groups, e.g., their attendance, achievements, and sex.

[6] As noted in the introduction, BHA uses the term "participant" to refer to individuals who are directly participating in resilience activities. Individuals who may benefit indirectly, e.g. community members who may receive free radio broadcasts, may be referred to as "indirect beneficiaries" or "non-participants."

- Community-based health workers, midwives, or other volunteers collect information directly from participants when they receive services or extract information from participant health records during home visits or at sites of growth monitoring, food distribution, and health or nutrition education, e.g., date of antenatal or postnatal care visits, quantities and types of commodities received, and date of receipt.
- Participants, community-based workers or volunteers, and frontline activity staff diaries that keep record of practices, productions, sells, input purchases et cetera. The data are typically noted in the diary or marked in a pictorial notebook. Program staff responsible for data collection verify data for reliability and validity from a randomly selected sample of participants. Once satisfied with the validity and reliability, activity staff transfer the data into data collection forms and either enter into a database or report to the next level depending on data flow process.
- Members of participant groups or communities record information about their activities and submit written reports to activity field supervisors monthly or quarterly. Sectoral supervisors verify the data for validity and reliability and extract and aggregate information from the various groups' reports, e.g., information from savings and loan groups about loans granted and repaid or the types of actions taken by community committees, to enter into a database or for submission to M&E staff.
- Activity specialists review community early warning or disaster risk reduction plans at regular intervals (e.g., twice annually) to assess whether they are complete, viable, and on schedule, and record a score, as outlined in the PIRS, for each community plan.

Participant-Based Surveys (PaBS): The second approach is to collect information from a sample of participants using surveys. There are two participant-based survey designs used for BHA programming—annual and routine—and they require different recall periods and frequencies of data collection.

Annual participant-based surveys consistently take place at the same time each year. The survey may take place at intervention sites or in homes. The timing might be near the end of the FY or it could relate to season. For example, annually, shortly after the planting season ends, the M&E team could systematically sample FFS participants to ask a set of questions related to plot sizes, the preparation of land, seed sources, and planting of specific crops. The data collectors would have a limited time window in which to collect and report information, using a standard questionnaire, and only from the selected sample participants.

Most BHA activities include numerous interventions, each of which benefits different individuals, households, and communities. Each monitoring indicator relates to participants involved in a specific set of interventions. **When designing an annual survey, a sampling frame must be identified for each indicator for which data will be collected.** BHA discourages the use of large household surveys to simultaneously collect information for indicators that relate to different intervention groups because of the difficulties in ensuring the accurate coverage of households from more than one sampling frame. This is especially difficult when a given indicator's sampling frame is small relative to the others.

BHA encourages the use of a PaBS **only** for those indicators for which data collection through routine monitoring is determined as not feasible or unreliable. When the partner chooses to conduct an annual PaBS to collect information from multiple participant groups, the AOR must approve the survey SOW prior to implementation of the survey and, when the survey will be done by an external party, prior to solicitation. Note that BHA does not consider Lot Quality Assurance Sampling (LQAS) an acceptable method to generate sample sizes for point in time estimates for monitoring performance indicator data. (LQAS may be acceptable in rare instances to monitor coverage or monitor prevalence rate of uptake of a certain practice if a threshold prevalence rate cutoff point is identified in advance.)

Routine participant-based surveys gather information at regular intervals during the year from a sample of intervention sites or from a subset of participants using a probability sampling method. To derive the annual figure, all of the data collected through the FY are analyzed as a whole. This approach is particularly useful for measuring knowledge, attitudes, or practices, which generally require more time for questions and for which questioning of all participants would be onerous. The sampling strategy must be well defined and rigorously followed to ensure that the data collected represent the entire participant group and ensure comparability from one year to the next. Examples of data collection using routine participant-based surveys include:

- Regular visits (e.g., monthly, quarterly, twice a year) by program/technical/M&E staff to carefully selected samples of intervention sites to observe and record information about the implementation process or to interview a systematically selected sample of participants about their experiences or current knowledge, attitudes, or practices. Collection times and sampling approaches (sites and participants) must be consistent across years.
- Surveys by M&E staff of carefully selected samples of members of targeted community groups (e.g., youth, farmers, or livestock holders) at regular intervals (e.g., monthly, quarterly) to determine exposure to, knowledge of, or reactions to messages broadcast generally in the communities (e.g., through radio, on billboards, or in community meetings) or perceptions about or use of community services or infrastructure supported by the activity.

The partner must submit an annual PaBS SOW and/or routine PaBS survey plan to the AOR for approval each year. The SOW(s) must include:

- Justification for using this type of survey instead of routine monitoring;
- The indicators to be measured and the intervention groups in each sampling frame;
- Justification of the sampling design, and how many from each intervention group will be included;
- Sample size estimation calculations;
- Methods;
- Data processing and analysis plan;
- Partner and/or external contractor responsibilities;
- Partner staff and/or external contractor qualifications; and
- Timeline and deliverables.

3.1.3 QUALITATIVE MONITORING AND QUALITATIVE STUDIES

A comprehensive monitoring system requires data that is both quantitative and qualitative. Qualitative methods and tools often complement quantitative indicator data and provide information about more complex phenomena that are crucial to activity success. The complex theories of change underlying BHA activities are influenced by social, political, and other contextual aspects of the broader operational environment, and they also rely on successful social and behavioral change among direct participants. Tracking, understanding, and responding to these difficult-to-quantify aspects of the activity can provide important, actionable information for programming.

In the context of BHA resilience activities, qualitative monitoring (QM) may be used for process monitoring such as quality of behavior change sessions or demonstration plots, outcome monitoring such as women's empowerment, context monitoring such as conflict dynamics, unintended consequences, magnitude of inclusion and exclusion errors, and secondary adoption. Qualitative studies (QS) may be used to answer discrete questions that arise during implementation, provide explanations for patterns in quantitative data, or inform specific strategies.

Routine **qualitative monitoring (QM)** collects data at consistent intervals. It is used to gather information that the activity staff know they need on a regular basis in order to complement the data

generated through quantitative measures, thereby facilitating a more holistic understanding of the situation. A QM approach may be particularly useful for process monitoring.

A **qualitative study (QS)** is a discrete study to better understand dynamics or phenomenon related to higher-level results or outcomes, and specifically in cases where qualitative data will be more informative or appropriate than quantitative data.

Potential Use of Qualitative Methods and Tools in Monitoring

Qualitative inquiry is very useful for process monitoring and attaining information related to outcome and impact level indicators. Areas where a QM or QS is particularly useful include:

- **Process monitoring:** Sub-optimal implementation quality could be a major factor for an activity not to achieve its objective. The activity could be designed based on proven theories, yet adoption of activity promoted practices could be very low because of a variety of reasons including poor quality behavior change sessions, not using adult learning techniques in learning sessions, and poor-quality demonstration plots. To identify these issues an activity must use qualitative monitoring tools such as observation of sessions and plots, discussions with the participants, interviewing front line staff to understand their level of knowledge, visiting participants' homes and plots to observe signs of adoptions.
- **Unexpected and unexplained achievements:** Quantitative indicators may suggest that progress toward a quantitative target is not on track (e.g., when progress against targets is unexpectedly low or high). Qualitative methods or tools could be used to understand the reasons behind this under- or over-performance. The information then can be used to tailor the implementation strategy either to improve performance or use it as a positive deviance to inform other interventions.
- **Outcome monitoring:** There are anticipated outcomes that are not easy to quantify, therefore, qualitative tools and methods are suitable to capture these outcomes. For example, peoples' agency, empowerment, gender equity in decision making, coping strategies, and changes in norms and attitudes.
- **Post distribution monitoring (PDM):** While quantitative methods work well to capture data to track the use of transfers, they are not as effective to understand the reasoning behind the decisions. Qualitative methods are well suited to understand the dynamics of targeting, feedback on protection issues, transactional costs, and waiting time. It may also be more appropriate in specific settings (e.g., school feeding programs, in politically sensitive contexts where randomized sampling is not feasible).
- **Unintended effects**: Qualitative data collection is well-suited to explore possible unintended consequences or unexpected outcomes of interventions that would be overlooked in routine quantitative monitoring.
- **Effects of layering, sequencing, clustering:** Qualitative methods can be used to explore the interaction among interventions and how activities are best sequenced and clustered to achieve desired outcomes (e.g., linking WASH with small garden interventions or VSLAs together with alternative livelihoods programs and value chains).
- **Monitoring Sustainability:** Market-based extension services are critical to support sustainability of outcomes. Implementing partners must develop market-based local service providers and gradually transition from direct delivery to market-based service provisioning. Transitioning to high quality and effective local level service provisioning requires monitoring of the market-based service provisioning including quality of services, demand for services, effectiveness of the extension services, and willingness to pay. Qualitative data may help to

illuminate participant intentions regarding long term sustainability and maintenance of interventions.

- **Secondary Adoption:** BHA expects a population-level change for key outcome indicators. It is important for an activity to monitor secondary adoption. Without secondary adoption, an activity may not be successful even if most direct participants adopt activity promoted practices. Qualitative methods and tools may help to get a sense of the magnitude of secondary adoption and understand why certain practices are adopted by neighbors and what could be done to further promote secondary adoption.

3.1.4 QUALITATIVE SAMPLING STRATEGY

Designing the most appropriate sampling strategy for qualitative inquiry is critical to ensuring the findings are robust. The sample needs to reflect and be tailored to the purpose of the study, the research questions, and methodology.

Sources of data depend on the purpose of the inquiry. While the most common source of data are activity participants and their household members, valuable insight could be gained from non-participating neighbors. Data may also be collected from relevant stakeholders, including community leaders, representatives of government and nongovernmental agencies and private sector actors, as well as project implementing staff.

For qualitative inquiries, sampling typically involves "non-probabilistic sampling" or the deliberate selection of research participants to ensure collection of rich, detailed information. Purposive sampling is often used for qualitative inquiry. Purposive sampling techniques include snowball and quota sampling: snowball sampling begins with selection of one or more individuals who are expected to have useful insight into the study questions, and then the participant refers the interviewers to similar individuals or groups who are subsequently contacted for an interview. Quota sampling attempts to reflect the overall characteristics of the community being studied. It begins with estimations of various strata or cohorts within the community (e.g., male/female, youth, elderly, ethnic groups, wealthy, poor). Participants for qualitative research are then selected in a manner that approximates these same characteristics. In some cases, qualitative studies use convenience or opportunistic sampling approaches to collect information from participants who are easily accessible or to take advantage of circumstances as they arise. Convenience sampling is not recommended unless it is not feasible to identify the sample or the population.

Determining Selection Criteria

Sample selection criteria need to be established to address the research question and capture a desirable range of variation across the operating context. Validity of qualitative information relies on thoughtful selection and adherence to these protocols. Steps to construct a purposive sample and establish selection criteria are:

1. **Define the target community/village/area:** This is analogous to determining the "sampling frame," the population of individuals from which study participants will be selected. A qualitative sampling frame is typically defined by a geographic area such as a community within the project's operational area.
2. **Identify inclusion and exclusion criteria for the sample:** Most often multiple, complementary selection criteria are used. For example, to investigate the impact of drought on food security "small-scale farmers" would be the target population and selection criteria such as gender, household (HH) type, type of agriculture, type of terrain farming, participation in complementary activities, would be chosen as inclusion criteria to inform analysis of constraints for particular groups.

3. **Determine the sample size:** The determination of sample size should be guided by the question to answer and the selection of qualitative tools (e.g., FGDs, KIIs). Qualitative inquiry does not require probabilistic sampling using a pre-determined level of precision and power to generalize the results to the population with a defined level of confidence. The sample size for qualitative inquiry should be large enough to achieve a minimum level of saturation. **Saturation** is loosely defined as the point in qualitative research when there is enough data to answer the question and subsequent interviews can only add marginal new information or themes. In general, the depth of data is often more important than the number of interviews.[7] Note, it is possible to reach saturation prematurely, if, for example, the sampling frame is too narrow, the methods are not eliciting rich information, or the interviewer is not able to move beyond surface level attainment of information.[8]
4. **Select the targeted number of sampling units:** The final step in purposive sampling entails actually selecting individuals (sampling units) for participation. Create a plan to recruit and select individuals that satisfy the selection criteria and meet the proposed sample size. Then formulate a plan for enlisting participants.

Photographs: Photography can be a useful monitoring tool when used thoughtfully. Program staff must ask and receive consent before photographing individuals, and receive parental permission and child assent to photograph children.

3.1.5 ANALYSIS AND INTERPRETATION OF QUALITATIVE DATA

Data analysis is the process of bringing order to data, summarizing the information in a meaningful way, and organizing the information into patterns, categories, and basic descriptive units. **Data interpretation** means attaching meaning and significance to the analysis, explaining descriptive patterns, and looking for relationships and linkages among descriptive units. The first step is to organize, describe, and categorize data (analysis), and the second step is to draw conclusions from that analysis (interpretation).

Qualitative Data Analysis

Qualitative information collected for monitoring purposes may be transferred into matrices or inputted into qualitative software, and then analyzed to identify patterns in responses. These approaches are briefly described below. Qualitative data should not be turned into quantitative data. It is important not to oversimplify qualitative data that is rich in detail and nuance.

The following are some basic steps for starting the analysis:

1. Develop categories based on indicators or key questions.
2. Assign qualitative data such as quotes, descriptions, or summaries to the appropriate category.
3. Estimate values by counting, for example, how many people responded a certain way or behaved a certain way, disaggregating by gender, age, and other groupings based on selection criteria; and
4. Use direct quotes or descriptions to support the values.
5. Cross-check responses from participants between data sources (i.e., FGDs, KIIs, household surveys, and others) to ensure reliability of the information and to identify differences in perception between social groups based on gender, socioeconomic status, or ethnicity.

[7] See: Turner, D. 2016. Reaching saturation point in qualitative research. Quirkos Blog: https://www.quirkos.com/blog/post/saturation-qualitative-research-guide

[8] Cohen, D. and B. Crabtree. 2006. Qualitative Research Guidelines. http://www.qualres.org/HomeSamp-3702.html

Matrix (manual) approach for analysis: The matrix approach is a low-tech and proven method of organizing both data entry and analysis of qualitative data. It only requires team competency in using a spreadsheet or a table. This approach ensures that all team members are recording information consistently and in a manner that directly responds to key research questions. Capturing qualitative data in matrices enables identification of important patterns in responses and specific contextual information that may help to explain quantitative or secondary data. Developing qualitative data matrices also allows responses from FGDs, KIIs, households, and others to be triangulated to determine whether information is reliable. Before analysis of qualitative information can begin, the data is aggregated by the designated research leader into qualitative data matrices by location and/or implementation area. During the analysis, completed matrices are useful tools to ensure that qualitative information can be concisely and coherently presented either on its own or integrated with quantitative survey results.

This approach will be appropriate for most BHA resilience, recovery, and preparedness activity qualitative inquiries for the following reasons:

- Ensures that qualitative teams are consistently collecting and recording findings, but does not require transcription from recordings.
- Encourages concise recording of detailed data according to important themes.
- Facilitates iterative preliminary data analysis during research team debrief sessions, to identify patterns, information gaps, and key lines of inquiry.
- Allows quick comparison of qualitative data between men, women, children, and between different geographic locations.
- Enables consistent analysis of qualitative data from multiple groups.
- Helps researchers identify unanticipated findings through a manual review and synthesis of data matrices.
- Does not require the purchase of software or special training in software applications.

Software (computer-assisted) approach for analysis: In some instances, partners may wish to use a qualitative software package to store, organize, analyze, and synthesize unstructured or qualitative data and/or secondary data such as articles or reports. Analysts upload data (for example, interview transcripts) into the software system, and then code the data by themes following the established analysis protocol. The researcher can use data visualization tools associated with software packages to present descriptive results. Challenges of using qualitative software for analysis include the time and effort required to transcribe and code data before the analysis begins.

Qualitative Data Interpretation

The next step in analysis is to explain patterns or relationships observed from the analysis and triangulation, and to draw conclusions based on the research questions. The QuIPS should clearly articulate how the data will be disaggregated, which should help inform how patterns in the data across those disaggregates should be interpreted. For example, if the data will be disaggregated by gender and geographic area, the data should be interpreted to tease out differences or commonalities across these two disaggregates (i.e., by gender and by geographic areas).

3.1.6 REPORTING QUALITATIVE MONITORING AND STUDIES

BHA does not require implementing partners to submit stand-alone deliverables generated from qualitative monitoring or qualitative studies that are conducted after the refinement phase. (Partners should refer to the award language or consult with the AOR for specific guidance on deliverables associated with formative research or pilot studies conducted during the refinement stage.) Partners should use insights generated from qualitative monitoring or studies carried out during the implementation stage to support adaptive management, learning, and planning. **Partners should**

include findings from qualitative studies and monitoring in the ARR narrative, quarterly reports, and, when relevant, the PREP. If partners wish to generate stand-alone reports or other materials, BHA strongly encourages partners to limit the number and page length of any such deliverables and/or to create short executive summaries in lieu of lengthy reports.

3.2 DATA COLLECTION TOOLS

For both quantitative and qualitative monitoring, data collection tools should be standardized, to the extent possible, to ensure consistency. Qualitative tools should be developed as open-ended guides to generate descriptive responses. Often multiple tools are used to gain a comprehensive understanding as well as to triangulate data generated from multiple sources. Qualitative tools should be adapted to effectively guide meaningful discussions with various stakeholders of interest. Reasons for non-standardized tools, e.g., if different tools are used by sub-partners or at different locations, must be explained in the Monitoring Strategy section. Draft tools may be attached to the M&E Plan and/or Annual Survey Scopes of Work.

Feed the Future Indicator Data Collection Guidance. Partners should refer to the Feed the Future Agricultural Indicator Handbook (2019) for guidance on how to collect data for the following indicators:

- AM 9 (EG.3.2-25): Number of hectares under improved management practices or technologies with USG assistance
- AM 15 (EG.3-10, -11, -12): Yield of targeted agricultural commodities among program participants with USG assistance
- AM 16 (EG.3.2-24): Number of individuals in the agriculture system who have applied improved management practices or technologies with USG assistance
- AM 33 (EG.3.2-26): Value of annual sales of farms and firms receiving USG assistance

3.3 DATA FLOW

To ensure accurate values, data quality and usage, all collection, recording, transfer, storage, aggregation, disaggregation, and other processing of data should follow standardized, well-documented procedures. The Monitoring Strategy section should include data flow diagrams, matrices, or summaries for individual or groups of raw data types to show the flow of data from the point of collection through the various offices or individuals where they are verified, aggregated, disaggregated, entered into electronic devices, and otherwise processed to derive the values for the indicators that are finally reported in the IPTT. The accompanying narrative should add details about the nature of processing accomplished at each point and the mode, frequency, and timing of movement of data between points.

The data flow should begin by describing the location, process, timing, and actors involved in collection. It should note whether the raw data are recorded on physical media, such as participant logs, or data collection forms that are used at intervention sites or field offices. The flow should continue by describing how these data (in raw or aggregated form) are transferred (digitally or in hard copy) in reports or data sets to a regional or national office, and with what regularity (e.g., monthly, quarterly, preferably more frequently than annually) for further processing—addressing what, when, where, and who—up to the point of reporting. The data flow diagram should identify which data are entered and maintained in monitoring databases and how information from the databases feeds into annual reporting. See Figure 6 and Table 3 for two different ways to present the data flow for the same indicator.

If the presentation of the data flow differs by partner within a consortium, the Monitoring Strategy section must clearly identify the flow for each partner.

Figure 6. Illustrative Data Flow Diagram

Indicator: Number of individuals who have received USG-supported short-term agricultural sector productivity or food security training

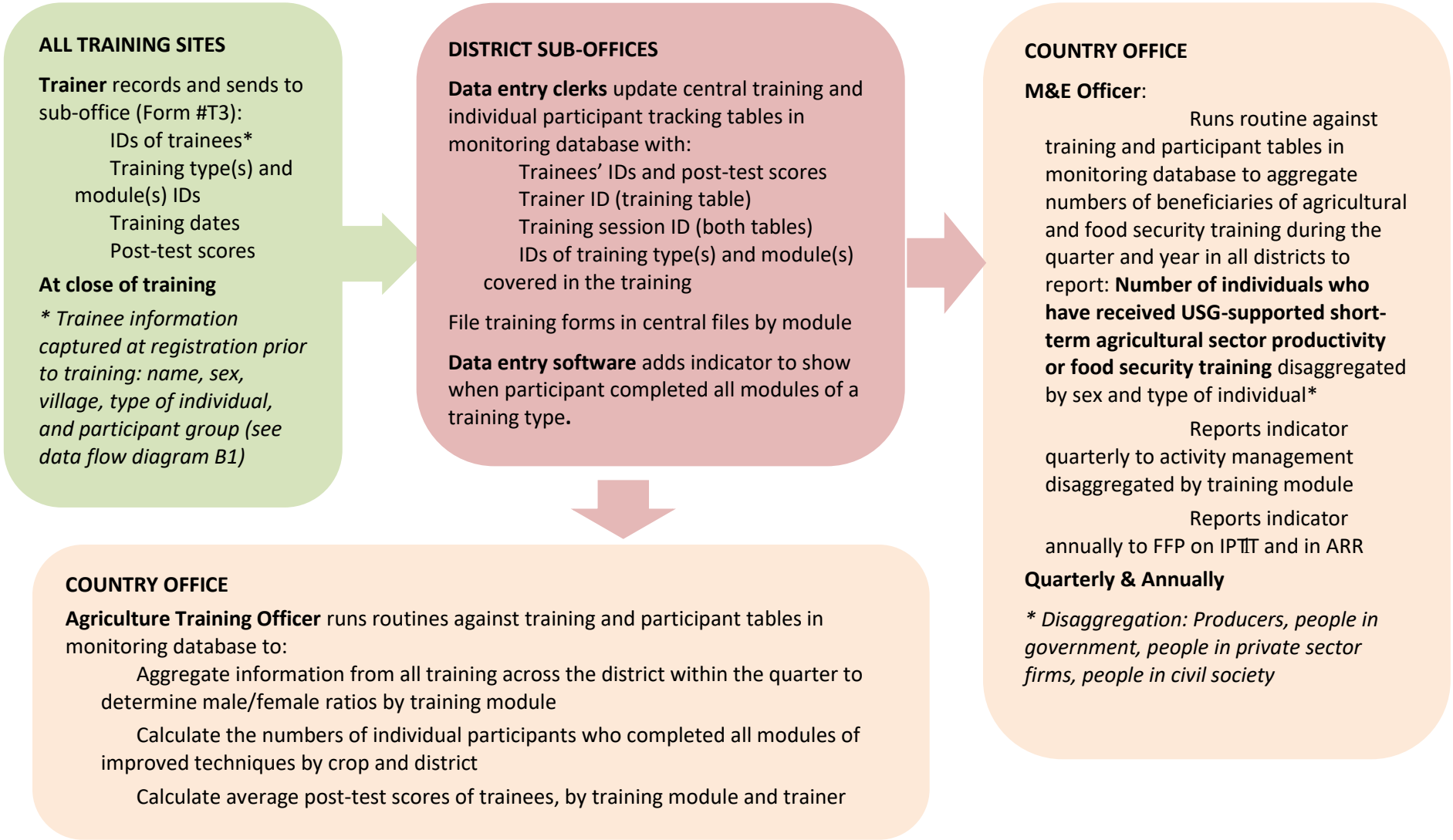


Table 4. Illustrative Data Flow Matrix

	WHO	WHAT	HOW	WHEN	WHERE
Collection	All trainers	► Fills Form #T3 (paper) with: • Trainees' participant IDs and post-test scores • Training session information: Trainer; Training dates; Type(s) and module(s) ► Sends completed form to sub-office via fax, scan, or courier	Paper form provided by Agricultural Training Officer; Email of filled form to sub-offices	Within 3 days of end of training session	Training sites/ trainers' offices
	Field staff / qualitative researcher	► Uses topical outline (TO) as a guideline for GDs, FGDs and KI interviews • Record raw data in notebooks ○ Include information on data collection event: location, number of men/ women/ youth; recorder, interviewer	First analysis done by data collector; emailed to Gender Advisor; Gender Advisor analyzes	Within 1 month of end of training session	Training sites during implementation
Entry	Sub-office data entry clerks	► Enters training table in monitoring database: • Trainees' participant IDs and post-test scores • Training session information: Trainer; Training dates; Type(s) and module(s) ► Paper Form #T3 filed in sub-office training file by training type and trainer	Training table data entry system	Within 1 week of end of each training	District sub-offices
Sto-	Database management staff	Maintains and safeguards monitoring database and data entry system		Indefinite	Cloud based
Processin	M&E Officer	• Aggregation of training data across districts • Disaggregation by sex, module • Disaggregation by sex, type of individual, and type of training, annually	Computing module T1 annually; Computing module T2 quarterly	Quarterly and annually	Country office
	Ag Training Officer	• Aggregation of training data by district • Calculation of: Male/female ratios by module by district; Numbers of participants who completed all modules of improved techniques by crop, sex, and district; and Average post-test scores by module and trainer	Computing module T3	Quarterly	Country office
Report	M&E Officer	Number of individuals who have received USG-supported short-term agricultural sector productivity or food security training	Manual transfer of numbers to MIS; include analysis in ARR	Annually	Country office

3.4 MANAGEMENT INFORMATION SYSTEM

BHA considers a Management Information System (MIS) as foundational to evidence-based programming. Often the terms "MIS" and "monitoring database" are used interchangeably; however, the MIS includes a monitoring database and a broader set of tools. BHA uses the term MIS to refer to a set of data collection, storage, and reporting software packages that partners use to monitor activities. BHA does not recommend or require any specific software package.

BHA recommends that all partners use an electronic MIS for BHA resilience activities. BHA may not accept annual results from partners who do not have a functioning MIS for monitoring data. BHA encourages partners to adopt mobile device-based monitoring data collection and avoid paper-based data collection unless absolutely necessary. The monitoring strategy submitted by partners must describe the structure and contents of an MIS.

MIS Software Selection: BHA recommends that partners select MIS software that are capable of managing relational databases (RDB), sometimes referred to as case-management software. These may include packages such as Kobo, CommCare, RedRose, and PowerBI; note that BHA does not recommend any particular set of software tools. A relational database (RDB) enables users to manage complex sets of datasets and preserve relationships/links between these datasets. BHA resilience activities typically implement multi sectoral interventions at individual, household, group and community levels. It is possible that several individuals and or households might participate in more than one sectoral intervention that leads to potential double counting during quarterly and annual reporting. The relationship-oriented design of RDBs enables partners to avoid double counting, generate summary tables with disaggregated data by geography, household, individual, group, sex, age, etc., and help create accurate sampling frames for annual surveys.

MIS Design: BHA recommends clearly defining the purpose and use of MIS for your activity before developing a tangible system. Ideally MIS should be designed simultaneously with IPTT and in collaboration with the activity program/technical teams. BHA encourages partners to wait until the first draft of IPTT before designing the relational database component of MIS. Partners should keep the copies of relevant PIRS, data flow maps/matrices, donor reporting requirements, and monitoring plans while designing MIS. Field testing an MIS and collecting early feedback from the users is a key to the high-quality system. Partners are encouraged to test early on the data collection form user design, data storage, management, and reporting processes and fine tune the undesirable aspects.

This section only provides key considerations for an activity MIS. Please see Annex IV for additional information and sources.

3.5 MONITORING CROSS-CUTTING TECHNICAL AREAS

In the Monitoring Strategy section of the M&E Plan, partners must describe how the activity's M&E system will monitor and measure the frequency and quality of efforts to ensure that all interventions consider the cross-cutting themes of gender integration, environmental protection, conflict sensitivity, community participation, and sustainability. Partners are encouraged to use both quantitative and qualitative methods to assess and monitor the degree to which these technical areas are integrated into activity operations. The Monitoring Strategy section should highlight all methods that the M&E system will employ beyond the use of BHA indicators in sections dedicated to these themes, including descriptions of the staff members involved, methods used, frequency, locations, and types of informants.

Monitoring Conflict Sensitivity and Community Participation

BHA expects partners to include members of the target communities as plans and decisions are made at all stages of the activity. When done well, inclusive involvement can help bring community members together cohesively while providing activity staff more-complete and -accurate input for decision making; on the other hand, non-inclusive engagement can reinforce power imbalances, divide communities, and more.

The M&E system should monitor the extent and quality of community participation and the benefits and challenges of involvement at each stage. It should record a partner's efforts to avoid harm and strengthen intra-community cohesion and cooperation. Partners should clearly describe how the activity's M&E system will monitor the frequency, extent, and quality of community input and activity-community interaction throughout the LOA. The strategy and tools to measure community participation must be context-specific, taking into consideration cultural norms and practices, activity Purposes, cross-cutting technical areas, and interventions. This monitoring will likely be primarily qualitative; however, in some instances it may be appropriate to generate quantitative data points from the qualitative information being collected.

Methods to measure community involvement might include:

- Proportions, rank orders and characteristics (e.g., leaders and common members, men and women, different socioeconomic groups, youth, disabled, direct participants and indirect beneficiaries) of community members involved in information exchanges, analyses of activity challenges and results, and input to and timely information about activity decisions, and perceived costs and benefits of their involvement.
- Community members' and activity staff members' perceptions about the quantity and quality of information exchanged between them.
- Community members' perceptions about the activity's responsiveness to the information they provide.
- Knowledge of the activity and community about one another.
- Frequency and nature of information received from and delivered to communities and degree or consistency of follow-up to/by the communities.

Participatory Monitoring

Activity participants and community members (or others being interviewed or observed through the process of monitoring) are not merely informants in a participatory monitoring system. Participants identify the information that interests them and important to them for discussion. They analyze that information, gain insights from mistakes and successes, learn from neighbors, and develop an action plan to overcome the challenges. Activity staff play a facilitation role and help participants document the information. The activity staff then copies the relevant information and use it as a source for their own performance monitoring.

Participatory monitoring provides insights into the programming environment and how interventions are perceived by the community. Participatory monitoring is a great tool for communities to track their progress throughout implementation as participants invest significant time and resources by engaging in the activities. This information can be useful both for activity management and to increase interest and ownership among participants.

The design of participatory monitoring should be simple, and the information gathered should be based on participants' needs and interest. For example, participants who plan to increase production through applying improved crop or livestock management practices may like to know whether they are getting

higher production; the participants who plan to increase access and consumption of nutritious vegetables through developing home gardens may want to know whether their access to and consumption of vegetables improved. Mothers who plan to increase meal frequency and dietary diversity for their infants and young children may want to know whether they are progressing. These sessions are also very useful for participants who faced challenges, as they can learn from participants who did it successfully. When these sessions are designed in a group setting, the information can immediately be used by partner field staff to develop an action plan to help the participants who face challenges, and to better understand the challenges experienced by the participants.

The M&E team can help to develop the system and tools and train activity staff to facilitate these sessions. In addition to the examples above, the participatory monitoring system can also be a good tool for the following:

- **Intervention selection, targeting, and implementation:** Relevance, acceptability, consistency, cultural appropriateness, effectiveness, consideration and use of community members' capacities.
- **Accountability and cross-checking:** Perceptions about staff interaction and responsiveness; accuracy of staff reports.
- **Learning and adapting:** What worked and what didn't? Why do some adopt certain behaviors while others don't? What are perceived benefits of interventions, Outputs, and Outcomes? Were there unanticipated results (positive and negative)?
- **Communication:** Knowledge of the interventions, roles of the implementing staff and USAID, information that is meaningful to and desired by the target population, accessibility of formats and media used.

Community members should also be viewed as consumers of monitoring data. Feedback loops that help stakeholders understand how their data and analyses are informing activity implementation can lead to stronger buy-in, which can result in higher quality data, more robust analysis, and improved adaptation. Partners are encouraged to share information collected through the activity as a way to maintain feedback loops and dialogue with the community throughout implementation, as well as to demonstrate transparency and accountability to the communities in which BHA works.

BHA does not require that partners incorporate participatory monitoring into their M&E Plans and does not recommend any particular methods or tools for participatory M&E. However, when partners do rely on community members to collect or analyze data (e.g., in qualitative monitoring efforts), the community members' roles and responsibilities should be clearly articulated in the QuIPS and Monitoring Strategy, as relevant and appropriate.

General Resources

- Feed the Future Agricultural Indicator Guide provides guidance on the collection and use of data for selected Feed the Future agricultural indicators.
- McAID developed by Save the Children, and I-SMART, developed by ACDI/VOCA, are examples of database systems for monitoring.
- An issue of Participatory and Learning Notes provides articles related to participatory monitoring and evaluation.
- The World Bank's Sleeping on Our Own Mats: An Introductory Guide to Community-Based Monitoring and Evaluation offers an approach for activities and tools.
- The Food and Agriculture Organization of the United Nations offers a short Training Module on Participatory Community M&E.

- The USAID Center for Development Information and Evaluation offers a brief on Conducting a Participatory Evaluation.
- TOPS FSN Network (TANGO International and BHA), Monitoring and Evaluation Facilitator's Guide. Module 2: Qualitative Tools and Analysis for M&E, 2015.

Qualitative Data Collection and Analysis Resources

- FHI360, Qualitative Research Methods: A Data Collector's Field Guide (2005)
- The Food Security Information Network (FSIN) technical brief Qualitative Data and Subjective Indicators for Resilience Measurement (2015) provides guidance and examples for qualitative approaches to resilience measurement.
- Michael Patton's Qualitative Research and Evaluation Methods (2015) and H. Russell Bernard's Research Methods in Anthropology: Qualitative and Quantitative Approaches (2006).
- The Governance and Social Development Resource Centre (DFID) Quantitative and Qualitative Research Methods in Impact Evaluation and Measuring Results, by Garbarino and Holland (2009) discusses approaches, challenges and opportunities for mixed methods design for impact evaluations.
- J-PAL's A Practical Guide to Measuring Women's and Girls' Empowerment is a resource for establishing an iterative qualitative-quantitative process to measure women's empowerment (and other challenging concepts quantitatively).
- The Tufts University Positive Deviance Initiative developed a useful and brief Basic Field Guide to the Positive Deviance Approach (2010).

3.6 DATA QUALITY ASSURANCE, MANAGEMENT, SAFEGUARDS, AND UTILIZATION

M&E Plans must describe how a partner will ensure data quality and protect data throughout the processes of collection, transfer, processing, reporting, and storage. The way data are routinely collected, checked, and managed helps ensure data quality. To ensure data quality, the M&E Plan must incorporate a system and plan to protect the integrity of data as they are handled and managed routinely, to include cross-checks to test validity and reliability, as well as a process for supervising and verifying the data to test its precision and integrity of values collected. In addition, internal Data Quality Assessments (DQAs) should periodically test the rigor and effectiveness of these processes for a few key indicators. Finally, the M&E Plan must describe the measures taken to safeguard the data from inappropriate access, use, or manipulation.

3.6.1 DATA QUALITY ASSURANCE

The M&E Plans for all BHA resilience activities must describe the routine measures partners will take to ensure the quality of quantitative and qualitative data collected and generated by their M&E systems. The Data Quality Assurance section of the plan should describe how the partner will ensure that data meet the five key attributes of high quality: validity, reliability, timeliness, precision, and integrity. It must describe the strategies incorporated into the routine monitoring to reduce biases and errors in measurement, transcription, and processing, and the procedures for verifying and validating the data that are collected or generated by the systems.

In addition to the PIRSs and QUIPSs, partners must describe measures to protect and verify data quality by ensuring:

1. Complete, detailed documentation of methods and protocols for every process related to data collection, cleaning, recording, aggregation, disaggregation, documentation, access, safeguard,

reporting, and storage, written in easily understood language and readily available to the collectors and processors at the sites of collection and/or processing.

2. Regular documentation of verification that all data collectors and processors are cross-checking and following the documented methods and protocols consistently. Cross-checking and verification methods include such measures as:

    - Visits by supervisors or M&E officers to a sample of farmers, mothers, or households to verify the information previously collected by community volunteers or partner staff.
    - Inclusion of photographs, video or audio recordings, or other evidence to allow others to verify observations, transcriptions, and interpretations by the collector.
    - Triangulation by asking the same thing in different ways or in different contexts.
    - Systematic review of collected data to compare values collected across time and location to flag outliers or reversals of trends that should be investigated.
    - Incorporation of reasonability checks and comparisons into data collection, entry, and processing software; double-keying of data in entry procedures; use of dropdowns and conditional entry fields; and developing filters, macros, and scripts to identify data outside reasonable parameters or data that contradict each other.
    - Adequate staffing by individuals who receive regular supervision, training, and support to build and maintain their capacity, and who are held accountable for good performance, considering clarity of task descriptions of the specific functions related to data collection, recording, and processing.
    - Adequate financial resources and logistical support to ensure timely performance, e.g., for travel, training, and procurement/reproduction of instruments and tools.

The partner must describe measures for ensuring data quality in detail in the M&E Plan. The partner may prepare a stand-alone section in the M&E Plan to describe all data quality assurance measures for all monitoring indicators. Alternatively, descriptions of some or all of the data quality assurance measures may be incorporated into indicator PIRSs or the data flow description in the Monitoring Strategy section of the M&E Plan. Similarly, descriptions of staff functions and capacity building to ensure data quality should be included in the M&E Staffing and Capacity Development Strategy section.

3.6.2 DATA QUALITY ASSESSMENT

Data Quality Assessments (DQA) are periodic reviews to assess how well the data quality assurance processes have ensured that indicators reported in the IPTT meet the five standards provided in USAID ADS 201.3.5.8: validity, reliability, timeliness, precision, and integrity. The purposes of a DQA are to identify factors that contribute to higher or lower quality and ways to improve quality.

Each activity DQA reviews the quality of selected monitoring indicators that are collected through **non-survey methods**. Annual survey indicators are not included because their quality must be verified via supervision and verification during the survey.

For each internal DQA, BHA recommends that a partner focus on 3–5 indicators for each year of implementation. Selection should be strategic, for example:

- Indicators that are complicated to measure;
- Indicators of suspect data quality;
- Indicators that are very important to support decision-making or make a determination of an activity's progress; and/or

- Indicators that represent distinct or new data flow processes or procedures.[9]

In the DQA process, reviewers reconstruct the flow of data for each of the selected indicators to verify the quality and potential sources of error at every stage, beginning from the initial point of collection and continuing through to the highest level of reporting and use. The legacy Office of Food for Peace (FFP) DQA webinar and MEASURE Evaluation Data Quality Assessment Methodology and Tools are useful resources for DQAs.

A DQA process may examine:

- M&E structure, functions, and capabilities;
- Indicator definitions and reporting guidelines;
- Data collection tools and reporting forms;
- Processes of data verification, aggregation, processing, management, storage, and safeguarding;
- Data use and dissemination practices; and/or
- Links with national reporting systems (where relevant).

In the initial submission of the activity M&E Plan, the partner must describe the timing and processes of internal DQAs planned to take place during the first 12 months of implementation or as agreed upon with BHA and Mission staff. Because interventions may not get under way for several months, it may be too soon to conduct a DQA within 12 months of the first submission. In this case, the first submission of the M&E Plan should explain that no DQA is scheduled during the first 12 months, outline basic criteria for selecting the indicators to be assessed in the first DQA, and estimate the approximate timing of the first DQA. With each PREP submission, a partner should describe all DQAs planned for the following 12 months or provide a strong justification for why no DQAs are planned.

The description of activity's DQAs must include:

- A list of indicators to be reviewed and justification for the selection;
- Time frame, i.e., timing and duration;
- Any particular focus of the review; and
- Who will participate in the DQA along with their roles and qualifications.

For each FY, reports of the DQAs completed during the year, including a description of the DQA, the assessment findings, and actions taken in response to the findings, must be uploaded to ART as part of the annual ARR.

During the LOA, the relevant USAID Mission may also conduct one or more DQAs for a selection of activity indicators. These typically examine only indicators used by the Mission for annual reporting. The Mission DQAs do not replace the need for partners to conduct DQAs to test the adequacy of their data quality assurance functions, and the USAID Mission DQAs are not part of the partner M&E Plan.

3.6.3 DATA MANAGEMENT AND SAFEGUARDS

The submitted M&E Plan must describe a partner's plans for protecting data from unintended change, misuse, loss, or destruction as it is collected and as it flows between and through the various sites of processing to its final storage location. This relates to data on paper, on other media, and in digital format. Any breach of privacy or inappropriate use of data can potentially result in negative unintended

[9] As the DQA assesses the effectiveness and integrity of the data collection, storage, aggregation, and reporting systems, a partner may want to classify the indicators according to similar data flow (e.g. those that follow the same processes of data collection, storage, aggregation, and reporting) and select indicators to represent different groups.

consequences, especially in contexts with conflict or internal divisions and tensions. Therefore, access to data for viewing, use, and modification must be restricted. The plan should also describe how and for how long the data will be preserved for future use, such as for evaluations, DQAs, or future studies. For activities with a consortium of partners, i.e., with both a prime awardee and sub-awardees, the M&E Plan must describe how data management will be coordinated across partners.

Examples of data management and safeguards include:

- Measures that will be taken to ensure and safeguard participant confidentiality and protect personal identity information contained in quantitative and qualitative data, both hard copy and digital files;
- Systems to store/maintain original data files/activity records: Where original data will be stored, how they will be protected, who can access them, how long the partner will retain them, and procedures and timeline for their destruction;
- Methods, frequency, and locations of file and database backups; including a designated staff member responsible for making backups; and
- Measures to prevent and detect unauthorized data access for data entry, editing, processing, or retrieval; virus protection of digital data; and security measures to protect the physical location of hard copies, databases, and data backups.

3.6.4 DATA UTILIZATION STRATEGY

What happens after data are collected is perhaps the most critical aspect of monitoring in the Program Cycle. For monitoring efforts to truly serve their purposes, plans for how the data will be used need to be considered across a variety of potential uses and audiences. Without the final steps of determining how the information will be used to inform programming, the time and resources devoted to collecting the data will have been wasted. Learning should be emphasized throughout the life of the award, and the data utilization strategy component of the M&E Plan helps ensure that BHA and the activity staff can learn from monitoring data and adjust interventions and approaches, as needed.

In addition to tracking whether performance targets are being met, data should be regularly reviewed to understand the quality and effect of interventions, and to assess the progress towards the hypothesized pathway of change. The data utilization strategy needs to describe how information gathered through monitoring and qualitative studies will be used to: 1) test the hypothesis used to develop the TOC and make adjustments is needed; 2) assess whether the implementation of interventions are high quality to affect the anticipated change; 3) use lessons learned to continue or adjust interventions, or implementation approaches; 4) inform implementation adjustments as circumstances change or learning evolves; and 5) reporting and storytelling of the activity's impact. Data should be utilized for TOC reviews, periodic review of monitoring data, dedicated 'pause and reflect' sessions, work planning, and reporting.

The data utilization strategy should detail the frequency with which learning, reflection, and reporting occur and identify the people or stakeholders involved, in alignment with the activity work plan and seasonality. Analyzing and using activity monitoring and qualitative data on a regular basis helps partners and BHA identify what is working and what needs adapting, especially taking into consideration the effect of changes in the operating environment or context.

3.7 M&E STAFFING AND CAPACITY DEVELOPMENT STRATEGY

It is essential that the M&E Plan demonstrate that the activity has adequate personnel with sufficient capacity to carry out all processes of data collection, processing, and reporting; as well as to ensure data quality assurance; and to manage and safeguard the data. The M&E Plan must include a M&E Staffing and Capacity Development Strategy. **This should identify all staff members—both M&E and non-**

M&E—as well as external actors who will contribute to data collection, processing, management, and reporting. The list should include more than the M&E staff members whose time is fully or largely dedicated to M&E. It should also include the field staff, sectoral specialists, staff members of partner organizations and government agencies, consultants, volunteers, and other members in participant communities who contribute to any monitoring function, and should identify the monitoring roles and responsibilities for each position and the percentage of time incumbents will devote to each function and responsibility.

The plan must include an organogram that graphically displays the lines of supervision and reporting among the identified actors and their M&E functions. In the case of partnerships or consortium-managed activities, the plan must identify the organization for each position in the descriptions and on the organogram.

Partners should present the M&E Staffing and Capacity Development Strategy in the award application. This strategy should include how staff will be trained to collect data, applying gender-equity, environmental protection, community engagement, and "do no harm" principles. The M&E budget should identify the costs related to capacity building.

Partners may use a variety of approaches and modalities for developing staff capacity, such as formal or on-the-job training, mentoring, distance learning, and rotations.

CHAPTER 4: EVALUATION PLAN

The partner M&E Plan must include an Evaluation Plan section that includes the following components:

- *Baseline Study:* the applicant/partner should articulate how they plan to support and participate in the baseline study preparation, and how the applicant/partner will use data and findings from the baseline study. Note that the final determination of the baseline and interim/final evaluation design, i.e. either a performance evaluation or impact evaluation, may not be final under after the award is made; applicants may refer to the baseline and interim/final evaluation in general terms if no specific methodology is provided to prospective applicants in the pre-award stage.
- *Midterm Evaluation (MTE) and Implementation Quality Review (IQR):* the partner should plan to manage the MTE process, so this section should outline the timing and high-level objectives of the MTE. This section should also include an option for the partner's participation in an Implementation Quality Review, which may be conducted either in addition to or in lieu of the MTE.
- *Interim/Final Evaluation:* the applicant/partner should articulate how they plan to support and participate in the interim/final evaluation preparation. Note that the final determination of the interim/final evaluation design, i.e. either a performance evaluation or impact evaluation, may not be final under after the award is made; applicants may refer to the baseline and interim/final evaluation in general terms if no specific methodology is provided to prospective applicants in the pre-award stage.

4.1 BASELINE STUDIES

A baseline study is required for all BHA resilience activities. The purpose of the baseline study is twofold: (1) to provide information to partner(s) about the activity's target population that can help improve the design and targeting of interventions; and (2) to serve as a reference point for values collected in the final evaluation. The partner's M&E Plan should include a section in the Evaluation Plan that explains how the partner will utilize data from the baseline study.

BHA manages the baseline study and a research/evaluation partner conducts the baseline study. A baseline study takes place during the first year of the award, before significant implementation of the activity has begun. The research/evaluation partner will use BHA's household survey instrument to collect data on and establish baseline values for current BHA BL indicators as well as additional, custom indicators as agreed upon and approved by the AOR. The BHA indicators are collected and tabulated as described in the *BHA Indicators Handbook Part I: Indicators for Baseline and Final Evaluation Surveys.*

Annex I provides a detailed overview of the baseline study methodology and requirements.

Baseline Study Methodology

BHA baseline studies are mixed-methods, comprising both a quantitative survey and a qualitative component. The qualitative component usually follows the quantitative data collection so the qualitative data can be used to interpret the quantitative data and generate a deeper, more nuanced understanding of the operating environment. At times, however, the qualitative inquiry may precede the quantitative data collection, such as to refine, validate, or contextualize the survey instrument.

The baseline study may be designed as part of a **pre-post performance evaluation design** to understand how an activity may be contributing to BHA-approved indicator targets, or an **impact evaluation design** to compare the activity target population against a comparison group to determine what could have happened in absence of the activity. If a randomized control trial or quasi-experimental design is used, the baseline survey may include a subset of additional indicators. The evaluation design

will be decided by BHA collaboration with the implementing partner, the research/evaluation partner, and the Mission.

Regardless of the number of BHA resilience activities in a country, the research/evaluation partner typically produces one baseline study report. Data is presented in aggregate for BHA's full implementation area, as well as stratified results by partner.

Partner Involvement in the Baseline Study

Partners play an important role in the planning, implementation, follow-up to the baseline study. Most importantly, partners are expected to participate in the **baseline planning workshop** hosted by BHA or the research/evaluation partner leading the baseline study. The purpose of this workshop is to contextualize the indicators, discuss the scope of the baseline study, timeline, and methodology, brief the research/evaluation partner on the political, social, and cultural norms and context in which the baseline data collection will be taking place so that "do no harm" principles inform its data collection approaches; contextualize select questions in the survey questionnaire; and finalize the indicator reference sheets for custom indicators.

Prior to the baseline planning workshop, partners should **identify and submit a PIRS for any custom indicators** they request to be included in baseline and interim/final evaluation surveys. If any indicator requires context-specific validation through qualitative inquiry, rapid qualitative assessments and analysis may be conducted prior to finalization of the survey instrument. In addition to the required BHA indicators, each partner may request the inclusion of a limited number of custom BL indicators. The inclusion of these custom indicators could result in minor differences in the questionnaires used in the different partners' activity areas. Careful consideration should be given to the selection of custom indicators as additional indicators add to data collection time and cost and may result in poor-quality data due to respondent fatigue.

In addition to participating in the baseline workshop, the partner should:

- Brief the research/evaluation partner on the cultural, political, and social contexts in which it will be collecting data to help it develop a survey design that will do no harm.
- Provide a list of the activity's implementation communities/villages so that the research/evaluation partner can develop the survey sampling frame.
- Recommend communities for qualitative inquiry.
- Observe enumerator and supervisor trainings.
- Introduce the research/evaluation partner to the communities prior to data collection.
- Participate in periodic conference calls to receive updates on data collection.
- Review and comment on the draft baseline study report.[10]
- Attend presentation(s) on the findings (in-country or via webinar).

The research/evaluation partner may consult with the implementing partners to identify and select sample communities for qualitative inquiry; however, the sample communities must conform with the sampling strategy for the qualitative inquiry. This may be particularly appropriate if the activity programs in culturally or geographically diverse areas. Similarly, a few of the required BHA BL indicators depend on the local context and award-specific interventions and need to be contextualized by each partner. For example, the indicator "Percent of producers who have applied targeted improved management practices or technologies" requires each partner to identify the management practices and/or

[10] While BHA encourages the partner to review and comment on the baseline report, the research/evaluation partner has the right to agree or disagree with any comments. An implementing partner may submit an addendum for the baseline study report to record points of disagreement or concern with the report contents.

technologies targeted by the activity. It is important to note that the management practices/ technologies targeted by the activity must be based on an analysis of growth potential, environmental considerations, and sustainability. Partners will also need to review the approved EMMP analysis to select techniques and practices that are relevant to the climate and environment.

To maintain the independence of the study, the implementing partner neither participates in the household sample selection nor observes any data collection. However, partner staff may participate in the baseline enumerator training.

How Partners Should Use Baseline Data

Baseline data can be particularly useful for re-examining the TOC to assess whether the TOC contains all necessary and sufficient interventions to achieve the anticipated food security and other outcomes. While the baseline study report will not make specific programming recommendations, it will provide the data and an opportunity for the partner to think through the implications of the evidence. For example, a baseline indicator may reveal that only 20% of the population in the implementation area has access to basic sanitation services whereas the partner previously thought this may have been closer to 50% at the application stage. With that current and context-specific information about the operating context, a partner can review its activity design to assess if incorporating additional or different interventions are essential to achieving the high-level results.

Toward the end of the first award year, BHA organizes a baseline data utilization workshop with the partner staff to help understand the situation and facilitate discussions about activity design and setting targets for key indicators. This workshop offers an opportunity for activity staff to work with the baseline data to understand the different pathways in the TOC. Following the workshop, partners will continue to refine their TOC and activity design by determining areas that need more or different interventions, and conversely, areas that may not need to be prioritized. The revised documents will be submitted to BHA for approval.

Reporting Baseline Values and Final Evaluation Targets

Partners will revise the TOC during the refinement phase and finalize the M&E Plan within a couple of months of the culmination workshop. The revised IPTT submitted with the revised M&E Plan must include baseline values for each baseline indicator and revised final evaluation targets based on the baseline values. Additionally, the partner must enter the baseline values into ART with the ARR submission after the end of the FY in which the values become available.

Resources

- BHA Indicators Handbook Part I: BHA Indicators for Baseline and Endline Surveys for Resilience Food Security Activities contains all current required indicators for collection in baseline and endline surveys.
- Supplement to Part I: BHA Baseline/Endline Questionnaire and Indicator Tabulations for Development Food Security Activities (Supplement) contains the model questionnaire

4.2 MIDTERM EVALUATION

4.2.1 MIDTERM EVALUATION OVERVIEW AND REQUIREMENTS

BHA requires an externally led Midterm Evaluation (MTE) for all resilience, recovery, and preparedness awards. The MTEs are primarily qualitative, process evaluations designed to gather evidence about the quality, effectiveness, and efficiency of activity implementation. MTEs gain firsthand knowledge about adoption of practices promoted by the activity, institutional arrangements, ground work to strengthen

local systems, progress towards sustainability of knowledge and technology transfers as well as maintenance of systems and structures build or repaired by the activity; identify and examine unexpected results, both positive and negative; assess the effectiveness of the targeting strategy; and develop recommendations for adjustments that will improve the activity's high-level results.

The MTE typically takes place approximately midway through the implementation period. The exact timing of the MTE will be agreed upon by the partner and BHA post-award.

The MTE will either be (1) managed by the partner and conducted by a third-party firm; (2) managed by BHA and conducted by a third-party firm; or (3) managed by BHA and conducted jointly by USAID staff, partner staff, and/or third-party consultants. The final decision of who will manage the MTE process will be made post-award.

MTE Managed by Partner: The MTE is usually managed by the partner, so applicants should budget for and plan to manage the MTE process at the application stage.

When the MTE is managed by the partner, the partner must comply with the following:

- Submit final SOW and estimated budget to the AOR within 15 months of the award. Partners should not disseminate the SOW or advertise for evaluators until the AOR has approved a final SOW, and should not contract evaluators before the AOR approves the MTE team members.
- The midterm evaluation team leader and sectoral experts must be approved by the AOR at least five (5)–six (6) months before data collection begins.
- The final midterm evaluation protocol must be approved by BHA before primary data collection begins.
- The final report must be submitted for BHA approval within three (3) months of completion of the field data collection. The draft MTE report should be submitted to the AOR, who will review and provide feedback for the MTE team before the report is finalized. The partner must submit the final MTE report to the DEC within 30 days of BHA approval and upload it to MIS as part of the ARR for the FY in which the report was approved.
- Within 30 days of AOR approval, the partner or contractor leading the evaluation will upload any quantitative data files to the Development Data Library, as appropriate, in accordance with USAID ADS 579: USAID Development Data and the award's standard provisions.
- Within 45 days of BHA's approval of the final report, or as indicated by the AOR, the partner will submit a Post-Midterm Evaluation Utilization and Action Plan.

Additional guidance on midterm evaluations is included in Annex II.

Midterm Evaluation Scope and Methodology

The MTE should focus on implementation processes, including quality of training, facilitation techniques, knowledge of the trainers and extension agents; activity management; monitoring; use of learning from various monitoring efforts and stand-alone studies; the integration of cross-cutting themes; and finally, the results toward anticipated activity results. The overarching purpose of the MTE is to assess the quality of implementation, for example, how successfully the activity implementation matches the proposed plans approved by BHA. The MTE also explains why delays, accelerations, or deviations from the strategy have occurred; identifies the strengths and challenges of how the interventions were implemented; and recommends adjustments to improve the effectiveness and acceptability of interventions in the targeted communities.

The MTE should also assess how various stakeholders (e.g., partner staff, participants, other community members, local experts, external partners) perceive the interventions and implementation methods to understand what they think is and is not working, what adjustments should be made, and why. This

should include documenting how well participant communities believe the chosen interventions match household and community priorities, including the degree to which the interventions do no harm and increase community cohesion. To identify strengths of and challenges to implementation, evaluators should compare implementation details and stakeholders' perceptions across purposively selected sites to assess the relationships of contextual and implementation factors with acceptance, relevance, efficiency, and the quality of implementation.

In addition, the MTE should seek evidence of high-level outcomes and sustainability and compare contexts, implementation, and outcomes across locations to identify factors that may be affecting the types and rates of change. To do this, evaluators should examine both intended and unintended changes. The evaluators should also compare the perceived results and unintended outcomes to the activity's TOC to validate, refute, or refine the pathways of change.

In countries with multiple awards from the same BHA solicitation, BHA encourages partners to collaborate and hire a single MTE team to implement a joint evaluation for all awards. A joint MTE would enable the evaluation team to work with multiple partners to identify opportunities for collaboration and coordination.

MTE Managed by BHA: If BHA manages the MTE, BHA will either hire a research/evaluation partner to conduct the MTE or use a joint MTE approach. If a joint approach is proposed, the MTE team may include technical staff from BHA, partner HQs or regional offices, and/or the Mission. When BHA leads a joint MTE, the tasks and timeline may be slightly different from those outlined in this section. BHA will discuss roles and responsibilities with the partner once it is determined that a joint MTE will occur.

Timing of the MTE

By the end of the first year of an award (preferably as part of the first PREP), BHA and the partner will establish when the MTE will take place, and whether it will be managed and contracted by BHA or the partner. The field-based data collection for the midterm evaluation should be determined based on when the evaluation team may directly observe implementation such as service delivery, trainings, or other interventions that involve direct participants. For example, an evaluation team member who will study the quality of asset development interventions should collect data while the associated labor is taking place (e.g., while workers are constructing an asset) or when the evaluator can observe the benefit of the assets. If Farmer Field Schools (FFS) are an important component of knowledge promotion, technology transfer, and behavior change, and the agriculture expert on the evaluation team should observe FFS sessions during the season when participants are learning and applying new techniques. There is no need for the MTE data collection to match the seasonal timing of the baseline study or final evaluation.

Preparation for the MTE should begin at least one (1) year before the results are needed. This is based on the following illustrative timeline:

- Drafting and approval of SOW: 12–16 weeks
- Procurement of evaluators: 12–16 weeks
- Secondary data review and MTE work plan preparation: 4–6 weeks
- Data collection: 4-8 weeks
- Validation workshop at the end of or shortly following data collection
- Data analysis and drafting of report: 6–8 weeks
- Feedback from USAID and finalization of MTE report: 4–8 weeks

When partners are managing the MTE process, BHA strongly recommends that partners begin the contracting process for the external team members as soon as possible to improve the likelihood that

high-quality consultants will be available to participate in the MTE (from planning to final MTE report submission) and that data collection may proceed on schedule.

When BHA is managing the MTE process, the timeline may change. The timeline will be discussed and agreed to by BHA, the Mission, and the partner(s) post-award.

Evaluation Team Composition

An MTE team consists of a team leader plus technical specialists; the team should include technical experts from all the activity's key sectors and cross-cutting areas. All evaluation team members must be external to the activity (i.e., not directly involved in the day to day management of the activity) and have qualitative research and/or field-based evaluation experience. While no member of the MTE team may have had any responsibility in the design or implementation of the specific activity under evaluation, a qualified individual affiliated with an implementing agency (e.g., a regional technical advisor, headquarters staff member, or a specialist working on an activity in another country) may participate in the MTE as an evaluation team member. Similarly, a USAID staff member who meets the competency criteria and has never had direct oversight responsibilities for the activity may participate as an evaluation team member. Only the MTE team leader must be both external to the activity and all organizations involved in activity implementation.

To avoid possible bias and potential disruption of activity implementation, the MTE team must not use partner staff as translators, enumerators, or supervisors. Further, partner staff, the BHA Officer responsible for oversight of the activity, the AOR, and any other USAID or partner staff with a direct stake in the activity may be involved in the MTE only as key informants and/or observers. These individuals may review and provide comments on data collection tools and instruments before they are finalized, and observe some of the MTE processes; however, these individuals may not collect primary data or participate in translation, analysis, or interpretation of these data.

When BHA leads a joint MTE, the approach, timing, and staffing, will be discussed and agreed upon by BHA, Mission, and partner staff.

Validation Workshop

The partner will be invited to participate in a validation workshop (or consultation) toward the end of or immediately after completion of the field-based data collection. This workshop will be led by the midterm evaluation team lead, and evaluation team members will participate. The purpose of the validation workshop is to provide partners the opportunity to respond to preliminary findings of the MTE team members and provide clarification or additional information that may be required to triangulate or validate those findings. Additionally, the workshop will serve as an opportunity to identify any additional data or information that the MTE team will need to collect (if there is still time and budget available to collect data). BHA, the partner, and the evaluation team lead will agree on the timing and scope of the workshop in advance, and these details should be clearly outlined in the final MTE scope of work.

Budget

Given that management of the MTE process will be decided post-award, applicants should plan to manage the MTE themselves and allocate a minimum of $200,000–$300,000 in the activity budget to cover costs associated with the MTE. This should include the cost of contracted MTE team members, international and local travel, and in-country lodging and per diem. Salary for technical specialists who are members of the partner organization may be charged for the days that they are directly involved with the evaluation. Other related costs that might be in the budget include expenditures for hiring local

personnel (e.g., drivers, translators, local technical experts), translating reports, and renting meeting rooms for presentations.

4.2.2 POST-MTE UTILIZATION AND ACTION PLAN

BHA expects mid-term evaluations to be used to guide the design and implementation of the activity following the evaluation. The validation workshop held toward the end of or immediately following the field-based data collection should enable the implementing partner to gain early insights into the MTE findings. In addition, within 45 days of either BHA's approval of the MTE report, or BHA and partner agreement on the recommendations from the joint MTE, the partner, in collaboration with the BHA Officer, must develop a plan of action to apply the MTE recommendations and submit it for AOR and USAID Mission approval.

Implementing partners should clearly prioritize the actions they will take based on recommendations in the MTE report (and/or from the validation workshop, as agreed upon in advance with the AOR). Prioritization of actions (e.g., "essential," "high," "moderate," "low," "impractical") should be based on the potential of each action to influence the activity's final Outcomes positively (e.g., "major," "significant," "minor," or "negligible") and available resources (time, material, money).

Partners are strongly encouraged to hold a post-MTE workshop with relevant staff, sub-awardees, and other stakeholders to inform the post-MTE planning process. This may include technical staff and leadership, BHA AORs, BHA and Mission technical staff, and Activity Managers. (Note that this is distinct from the midterm evaluation validation workshop, which serves as an opportunity for implementing partner staff to respond directly to the findings from the evaluation team members.) A post-MTE workshop, by contrast, would focus on what actionable steps the partner plans to take to refine the design, implementation approaches, improve quality of implementation, targeting strategy, or other concrete steps related to implementation.

Table 5. Template for Post-MTE Utilization and Action Plan

Activity:							
Country:		**Partner (Prime):**				**FY:**	
Recommendation	**Priority level**	**Resources required**	**Influence**	**Activity**	**Time frame**	**Responsible party**	**Indicator**
Participants for Purpose 2 (P2) agriculture and natural resource management interventions should be selected from Purpose 3 (P3) nutrition participant households.	High	$$	Major	Agree on revision to targeting criteria Modify written protocols and train staff to use the revised criteria Select all future P2 participants using the revised targeting criteria	By xx/xxxx	Deputy Chief of Party P2 & P3 Leads P2 & P3 field supervisors	Percentage of new P2 participants that are also P3 participants.

Once BHA and the partner agree to follow-up actions, they should establish a timeframe and a means of measuring progress and achievement for each action and assign responsibility for each action. The partner, using the agreed means of measuring progress, must describe progress in subsequent ARRs and, if/when relevant, PREPs.

Resources

- Ethical guidelines for an evaluation team are outlined in the American Evaluation Association's Guiding Principles for Evaluators.
- USAID Utilization and Learning from Evaluations.

4.3 IMPLEMENTATION QUALITY REVIEW

As noted in Chapter 1, BHA may decide to conduct an Implementation Quality Review (IQR) in addition to or in lieu of a midterm evaluation. This will be decided in consultation with the USAID Mission, partner(s), and BHA post-award. The IQR will be narrower in scope than the midterm evaluation and will not require a detailed report. BHA and the partner will agree upon the questions and methods that will be used for the review, as well as the timeline. The IQR may be conducted by: (1) an internal team led by BHA M&E team with the participation from BHA technical team members, partners' headquarters and/or regionally-based technical specialists, and technical staff from USAID Mission; (2) an external firm; or (3) a hybrid approach including both external consultants and USAID and partners' technical staff.

Similar to a midterm evaluation, the IQR may involve a validation workshop (to be organized by the IQR team lead) and a post-action planning workshop (to be organized by the implementing partner). BHA may request that the implementing partner create a post-IQR utilization and action plan (using the Post-MTE Utilization and Action Plan as a template).

4.4 INTERIM/FINAL EVALUATION

A final evaluation is required for all BHA resilience activities to assess performance against stated objectives and approved targets, except in rare circumstances when BHA may waive the requirement for a final evaluation. Final evaluations assess activity performance or impact; measure development outcomes; and inform the design of future activities with similar objectives and/or will be in the same or similar geographical area.

For R&I activities, BHA will manage an endline survey (i.e., the quantitative data collection process) that will be used for the interim/final evaluation (i.e., the entire evaluation process and report). The interim/final evaluation will be conducted in the fourth (or penultimate) year of the activity to help inform the decision to grant a cost extension for high-performing activities. If the activity is not extended, the interim evaluation will serve as the final evaluation. If the activity receives a performance-based cost extension, BHA will conduct a final evaluation at the end of the extension phase, and the evaluation conducted in the fourth year will be referred to as the "interim evaluation." The interim/final evaluation will take place during the penultimate year so results may be used to help inform decision-making about a possible cost extension for high-performing activities. Cost extensions will be decided using a variety of information sources and will be contingent on available funding and other factors that may be unrelated to activity performance.

Chapter 4 and Annex III provides additional information on the methodology and scope of interim/final evaluations that use the pre-post evaluation design. BHA will provide applicants and/or partners with additional information on the methodology and scope of impact evaluation designs on a case by case basis, when appropriate.

Evaluation Design

The interim/final evaluation will use a mixed method approach and will include both a quantitative and qualitative component. The qualitative component will address sustainability; local systems and institutional capacity strengthening activities; gender equity in decision making; social accountability and governance; and other relevant topics.

If a pre-post final performance evaluation design was used at baseline, then the endline survey will collect endline values using the same population-level sampling frame as the baseline. For a performance evaluation, the evaluation team will work closely with the partner to get a list of target communities and finalize the indicators.

If an experimental or quasi-experimental design was used to establish baseline values, the baseline and endline surveys may be either population- or participant-based. When an impact evaluation design is used, the evaluation team will closely work with the partner during targeting and the selection of communities and households from the beginning of the implementation period. In these cases, the partner will be asked to identify approximately twice as many communities as they plan to target so that the comparison communities can be selected to set aside. When a randomized control trial design is used, the evaluation team will randomly allocate the selected households or communities to either treatment (program implementation) or control (no implementation). Partners are expected to adhere closely to the randomization to prevent spillover effects from intervention to non-intervention communities. The evaluation team will collaborate closely with the partners throughout implementation to monitor adherence and address any issues that may arise during the implementation period.

Timeline

A final evaluation will take in the penultimate year during the same season as the baseline survey to ensure comparability of indicators that are sensitive to seasonality. The exact timing of the final evaluation will be decided by BHA post award.

Role of Partner

To avoid potential conflicts of interest, the partner will not participate in household selection or data collection. However, partners are encouraged to be engaged throughout the evaluation process, including during the review of the qualitative tools; identify the techniques and practices promoted by the activity, enumerator and supervisor training; introduce the enumerators to the village leaders; conduct in-country briefings; and review the draft report. Partners may be asked to do the following:

- Provide a list of the activity's actual implementation communities the survey sampling frame and study sites can be generated;
- Review the endline study protocol;
- Brief the research/evaluation partner on the political, social, and cultural norms and context in which the endline data collection will take place so that "do no harm" principles inform its data collection approaches;
- Observe enumerator and supervisor trainings;
- Introduce the research/evaluation partner to the communities prior to data collection;
- Participate in periodic conference calls to receive updates on data collection; and
- Review and comment on the draft report; a partner may submit an addendum to the evaluation report to document points of disagreement or concern with the report's contents.[11]

[11] While BHA encourages the implementing partner to review and comment on the evaluation report, the research/evaluation partner has the right to agree or disagree with any comments. The implementing partner may submit an addendum for the report to record points of disagreement or concern with the report contents.

Finally, partners should develop a strategy for how they will share or disseminate findings from the final evaluation both internally to their organization and externally with the wider community of stakeholders. This should be included in the partner's M&E Plan.

Partners must upload the final report and record interim/final evaluation indicator values as part of the ARR data entry for the FY during which the endline survey was completed.

ANNEX 1: OVERVIEW OF BASELINE STUDY

This annex provides a general overview of the baseline study methodology and requirements. It is intended to: (1) inform the development of a scope of work for the research/evaluation partner conducting the baseline study; and (2) to help BHA resilience, recovery, and preparedness activity implementing partners better understand the purpose, scope, and methods used for the baseline study and inform how baseline data should be used.

Note that this annex primarily refers to the baseline study methodology used for pre-post final performance evaluation designs. Some of the basic principles and considerations included here may be helpful to inform baseline data collection for an experimental or quasi-experimental baseline study.

Overview

BHA will manage an external research/evaluation partner to conduct the baseline evaluation. To reduce the risk of bias in the baseline survey, no firms or individuals involved in the design of the activity may participate in the baseline survey.

Before commencing the quantitative survey, the research/evaluation partner submits to BHA a survey design protocol that includes:

- A sampling plan, including sample design, sampling frame, sample size calculation, level of statistical precision and power, and respondent selection procedures;
- Training and field manuals for supervisors, enumerators, and anthropometry;
- Plans for supervisor, enumerator, anthropometry, and data entry training;
- A questionnaire translation, back-translation, pretesting/piloting plan; and
- Data analysis and treatment plan, including estimation procedures (sample weighting and other adjustments) and indicator tabulation and subgroup analysis.

Implementing partners will brief the baseline study research/evaluation partner on the cultural, political, and social contexts in which it will be collecting data to help it develop a survey design that will do no harm. The research/evaluation partner should also demonstrate how "do no harm" principles apply within the baseline study context. (The research/evaluation partner should also summarize in the baseline study report the steps it took to ensure the conflict sensitivity of its approaches, which will help maintain consistency during the endline data collection process.)

The baseline study research/evaluation partner must submit to BHA the survey design for approval to the AOR. The survey may commence only after BHA has approved the final survey design.

After the study is complete, the baseline study research/evaluation partner will submit a draft study report to the awardee for comment by the awardee and BHA. After responding to the comments, the research/evaluation partner will submit the final report for approval by the awardee and BHA. Once approved by the AOR, the research/evaluation partner must submit the report, supporting documents,

and all related data sets to the awardee in time for submission to BHA and the DEC and DDL within 30 days.

Baseline Study Statement of Work

The baseline study should include a comprehensive statement of work which should include, at a minimum, the following sections:

I. Introduction
II. Purpose and Objectives of the Baseline Study
III. Indicators for Collection
IV. Baseline Study Design and Methodology
V. Deliverables, Timeline, and Report Outline
VI. Contractor Responsibilities/Tasks
VII. Contractor/Firm Qualifications
VIII. Team Composition and Qualifications
IX. Ethical Guidelines\

Baseline Survey Design

BHA most often requires a baseline survey for a simple, pre-post evaluation design (BHA may use an experimental or quasi-experimental design in some cases). A pre-post evaluation design allows detection of statistically significant changes in Outcome and impact indicators by measuring them before (at baseline) and after (at endline) activity implementation.

The baseline survey must be conducted in such a way that a comparable survey is feasible as part of the activity's interim/final evaluation that will be conducted in the fourth (or penultimate) year of the activity. To ensure comparability, baseline and endline surveys should:

- Be implemented at the same time of year; and
- Use the same questionnaire and collect the same indicators in the same way.

Factors that can compromise the comparability between pre-activity and post-activity results include:

- Changes in activity coverage area over activity lifetime;
- Changes in indicators or indicator definitions between baseline and endline; and
- Inadequate sample size at baseline or endline or both.

Baseline Survey Sampling

Sample sizes for the baseline and endline surveys do not have to be identical. However, the sampling designs must be the same, and the samples must represent the same population for the surveys to be comparable.

The sample size should be sufficient to detect a certain percentage point difference between the baseline and endline values. The indicator used to determine the sample size will be decided upon by BHA and the research/evaluation partner. Sample sizes will be derived using the following parameters:

- 95% confidence level for one-tailed test;
- 80% power;
- Design effect of 2; and
- The sample should be appropriately inflated to account for nonresponse among sampled households.

To draw a representative sample, the research/evaluation partner must use probability-based methods, where every unit in the population has a known, non-zero probability of selection and where selections are made using well-established random mechanisms. Sampling frames should ideally include every cluster (e.g., village or community) in the BHA activity implementation area.

To reduce logistical efforts and cost, BHA recommends a baseline survey that uses multistage, cluster sampling. Recommended sampling stages are:

- First stage: random selection of clusters (e.g., villages, communities, or enumeration areas) from all clusters in the sampling frame.
- Second stage: systematic selection of dwellings within clusters after canvassing and listing all dwellings in the sampled clusters. Methods that do not qualify as probability-based methods, such as "random walk," are discouraged.
- Third stage: random selection of a single household within each selected dwelling.

BHA advises against a fourth stage selection of individuals within households. Data should be collected for all eligible individuals in the selected household for indicators related to individuals.

Qualitative Components

The baseline study will always include a qualitative component; this may include qualitative inquiry that is conducted prior to the quantitative data collection (e.g., to refine the survey questionnaire) or during/after the quantitative data collection (to inform or contextualize the interpretation of the quantitative findings). As with the quantitative methods, the baseline study scope of work will outline the qualitative methods that will be used, including the following:

- Data collection methods and design;
- Sampling methods, including inclusion and exclusion criteria; and
- Training, data management, and analysis and reporting plan (including data collection, transcriptions, translations, coding and analysis, and quality assurance and control).

Ethical Guidelines

The baseline scope of work must clearly state that every member of the evaluation team must adhere to ethical guidelines as outlined in the American Evaluation Association's Guiding Principles for Evaluators. A summary of these guidelines is provided below.

1. Systematic inquiry: Evaluators conduct systematic, data-based inquiries.
2. Competence: The evaluation team possesses the education, abilities, skills, and experience appropriate to undertake the tasks proposed in the evaluation. Evaluators practice within the limits of their professional training and competence, and decline to conduct evaluations that fall substantially outside those limits. The evaluation team collectively demonstrates cultural competence.
3. Integrity/honesty: Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.
4. Respect for people: Evaluators respect the security, dignity, and self-worth of respondents, activity participants, clients, and other evaluation stakeholders. Evaluators regard informed consent for participation in evaluations and inform participants and clients about the scope and limits of confidentiality.

5. Responsibilities for general and public welfare: Evaluators articulate and take into account the diversity of general and public interests and values that may be related to the evaluation.

Resources

- Sampling guidance for baseline and final performance evaluation surveys is provided in the Sampling Guide (with 2012 Addendum).
- The USAID Evaluation Policy provides information on the purposes of evaluation; the types of evaluations that are required and recommended; and the approach for conducting, disseminating, and using evaluations.
- USAID ADS 579: USAID Development Data provides guidance for complying with the requirement for all quantitative data to be stored in a central database (Frequently asked questions, USAID Open Data Policy).
- Ethical guidelines for an evaluation team are outlined in the American Evaluation Association's Guiding Principles for Evaluators.

ANNEX II: MIDTERM EVALUATION OVERVIEW

This annex provides a general overview of the midterm evaluation methodology and requirements. It is intended to (1) help implementing partners develop a MTE scope of work; and (2) provide additional guidance on the MTE methods and process for other technical staff who may participate in a midterm evaluation as an evaluator or observer.

As noted above, BHA may decide to conduct an Implementation Quality Review (IQR) in addition to or in lieu of a midterm evaluation. This annex can be used to inform preparation for an IQR, however, any deliverables or timelines for the IQR should be agreed upon in advance with the AOR.

Overview

Midterm evaluations are process evaluations and explore the quality of implementation including social and behavior change related interventions, how well the activity is following implementation plans and meeting targets; the acceptability of the methods employed to the participant population; and signs of changes that participants associate with activity interventions. The primary aim of a process evaluation is to assess the process and the quality of service delivery and how well the delivery of an intervention matches the original, approved plans and to identify factors that contribute to greater or lesser efficiency and quality of Outputs and to a greater or lesser acceptance of the interventions by targeted communities. The secondary purpose is to examine evidence of early changes in the target communities, positive and negative, intended and unintended, and compare them to the changes anticipated by the TOC and Logframe. The evaluation also should seek to identify the factors in the activity's implementation or context that appear to promote or impede the changes.

The MTE may be managed either by the partner or BHA. If managed by the partner, the partner must develop a MTE scope of work (SOW) and submit to BHA. If managed by BHA, the research/evaluation partner will collaborate with BHA and the implementing partner to create the SOW.

Midterm Evaluation Scope of Work

The process of developing the SOW should be collaborative and involve the relevant USAID Mission staff, BHA M&E team member, and BHA technical experts to collect input for objectives; key evaluation questions; and special requirements for the evaluation team, content, and timing. The scope of work may include the following components:

I. Introduction
II. Objectives & Evaluation Questions
III. Methods
IV. Research/Evaluation Team Responsibilities
V. MTE Team Composition, Qualifications, and Roles
VI. Implementing Partner Roles & Responsibilities
VII. Ethical Guidelines
VIII. Intellectual Property

Introduction. The introduction should outline the activity Goal, Purposes, and cross-cutting technical areas, and briefly describe the implementation setting, interventions, and factors that have significantly affected implementation so far. It should highlight new or unconventional approaches used in the activity's implementation and any concerns or issues deemed important to examine either by the partner or USAID. Because the MTE focuses on processes, the introduction should also include a description of the activity's management and operations, including M&E.

Objectives and Evaluation Questions. The SOW should clearly articulate the overall objectives of the evaluation (i.e. the "big picture" reasons behind the evaluation) and the specific evaluation questions that will be asked as part of the study. The MTE objectives may include, for example:

- *Objective 1: Soundness of Activity Design*. Assess the overall design strategy of the activity in terms of its relevance for addressing food insecurity with targeted impact groups, taking into account contextual changes that may have occurred since the activity began implementation. This will entail reviewing the strategies that ensure that the target groups are reached by the activity, reviewing the theories of change, and assessing the hypotheses, risks, and assumptions made during the design of the activity.
    - Evaluation Questions: What are the strengths, challenges, and lessons learned of the activity management/ implementation so far? Are the assumptions made in the theory of change that informs program design still valid? What changes have occurred in the context since the activity began? How have the activity strategies been designed to put in place the elements needed to contribute to higher-order social change? How should the activity's theories of change and results frameworks be refined or modified?

- *Objective 2a. Quality of Implementation.* Assess the quality of project inputs, implementation and outputs to identify factors that enhance or detract from the efficiency, quality, acceptability, and effectiveness of the interventions implemented and the likelihood that they will contribute to sustained achievement of activity' goals.
    - Evaluation Questions: In each technical sector, to what extent has the activity adhered to the initial technical standards, approach, implementation plan, outputs, and participant targets included in the initial technical narrative? What are the strengths and challenges to the activity inputs, implementation of interventions and processes, the quality of outputs and the sustainability of the outcomes achieved? How have problems and deterrents been managed?

- *Objective 2b: Coordination and Collaboration.* To expand on the overall quality of implementation objective: assess the level and effectiveness of coordination and collaboration with external organizations that are critical to high-quality implementation. This includes actors that provide complementary services necessary to achieve the outcomes, actors that will provide essential services to sustain the outcomes after the end of the activity, actors that influence participants' access to goods and services, and organizations that promote or impede an "enabling environment".
    - Evaluation Questions: What has been the effect of the various collaborative relationships cultivated by the activity toward enhancing the effectiveness or efficiency in use of resources? How effectively has the activity taken advantage of the other USG and non-USG investments in the same space to achieve cumulative impact? How aligned are the strategies of the activity toward the development strategies of USAID and the host country Government? What changes can be made in these collaborative relationships to further enhance effectiveness and efficiency?

- *Objective 2c: M&E and Adaptive Management.* To expand on the overall quality of implementation objective: review systems for capturing and documenting lessons learned and assess the extent to which they are used in activity implementation and refining activity design, including feedback from the perspective of stakeholders and participants. Assess processes to use evidence including baseline results and monitoring data for adjusting project strategies. Assess how well the activity is seeking out, testing and adapting new ideas and approaches to enhance activities' effectiveness or efficiency.

- Evaluation Questions: How have activity management and technical specialists used data to inform programmatic decisions, referral and follow up? What processes have been instituted to improve data collection and data quality? How has the activity improved effectiveness or efficiency as a result of new ideas or approaches brought into the activities? How is information generated by the activity used to inform decision-making? How can this be made more effective?

- *Objective 3: Activity Results & Evidence of Sustainability.* Present thorough quantitative data and qualitative information, and evidence of changes (intended and unintended outcomes) associated with activity interventions and outputs; assess how well the observed changes support the theories of change and logic of the LogFrame, and identify factors (both internal and external) in the implementation or context that impede or promote the achievement of targeted results. Then determine the extent to which outcomes, systems, and services are designed and being implemented to continue after the activity ends, and assess progress made on implementing sustainability strategies. What interventions are being implemented to ensure that the service providers will: (1) have continuous access to required resources, (2) receive capacity strengthening support, (3) create demand and influencing the motivations of the participants and service providers, (4) establish and strengthen critical linkages necessary to sustain resources, and (5) sustain capacities?
    - Evaluation Questions: What changes do community members and other stakeholders associate with activity interventions? Are there signs of early outcomes? Which factors appear to promote the apparent changes, and which have deterred intended changes? Has the activity developed and implemented sustainability strategies? To what extent are government officials, formal and informal local leaders (whose support and understanding will be critical for continuing program initiatives once the project has ended) involved in activity interventions and included in ongoing program discussions?

- *Objective 4: Gender and Inclusivity.* Relative to the major cross-cutting themes in the activity, determine the appropriateness and effectiveness of support for gender equity in terms of access to, participation in, and benefit from activity interventions. Assess the extent to which activity interventions target youth, support greater capacities for local governance and address sources of environmental risk.
    - Evaluation Questions. How effective are program design and implementation mechanisms in addressing the cross-cutting issues of gender, governance, the environment and targeting of youth? What (if any) challenges have projects encountered in these areas that may not have been anticipated in the project design, and how have the projects responded? In what ways is the project changing roles, relationships, communication and decision-making dynamics among women and men, young and old, in relationship to food security at the household and community levels? How were the findings and recommendations of the Year 1 gender analysis considered in the program strategy and project activities? Have gender gaps and related concerns been addressed adequately? Is the project drawing on the potential of women, men, boys, and girls as much as possible?

Additional objectives, specific to the activity, may be added to address questions or concerns of the partners or USAID. Alternatively, the gender and inclusivity objective may be incorporated throughout the other objectives and addressed through specific evaluation questions under each objective.

Either in this section or as an attachment, include a matrix that identifies detailed areas of focus for each question, aspects to consider within each focus area, and illustrative methods for investigation. Table A1 provides an illustrative example of such a matrix. This matrix should be tailored to the details and

concerns of each activity. The partner and USAID should use this matrix to clearly communicate to the evaluators the general and specific features and components of activity implementation that should be investigated and the level of detail that is desired. It should highlight areas of concern and optimism, as well as new or different approaches that were implemented in the activity.

Methods. This section should clearly articulate the overall evaluation design, including: the evaluation design; participant selection; methods for data collection and analysis; and any other key considerations.

- *Design.* The MTE should include qualitative, primary data collection and analysis; and secondary analysis of existing quantitative data (e.g., activity monitoring data). In general BHA will not support costs for administration of a quantitative survey as part of MTE primary data collection. Under exceptional circumstances, a partner may request approval from the AOR and responsible BHA M&E team member for such a survey, but the request must include compelling justification for the expenditure.
- *Participant Selection.* The methods section should clearly articulate which key stakeholder groups, e.g., participants, other implementing partners, government partners, and other external collaborators, who should be included among MTE informants. The MTE informants should include people representing the different population subgroups to ensure that all population subgroups are benefiting either directly or indirectly from an activity appropriately and to capture unintended Outcomes—positive and negative—for different population subgroups. It may be important to capture non-participants perspectives, for example, to gather information necessary to assess intervention coverage of intended participants (looking for evidence of both inclusion and exclusion errors) and to get outsiders' impressions of the interventions' implementation and relevance.
- *Data Collection and Analysis.* This section should describe how primary (qualitative) data will be collected and how secondary (usually quantitative) data will be collected. This may include a reference list and descriptions of secondary data resources that the partner will provide the MTE team (found in the "activity responsibilities" section of the SOW) and indicate how far in advance of primary data collection the resources will be available to the team. These resources should include documentation describing the activity's M&E processes, including monitoring databases; all reports from formative research, gender analysis, barrier analyses, and other special studies; and the data sets, analyses, and results of participant-based surveys.
- *Other Key Considerations.* This section may include any additional information related to the proposed methods, such as expectations and requirements for in-country review and validation of the evidence collected by the team, or an overview of any factors that may hinder investigation of certain topics due to contextual circumstances.

Research/Evaluation Team Responsibilities.

- *Deliverables.* This section should itemize all deliverables required from the evaluation team, including any specific requirements for the content, format, or length of the deliverables. Illustrative deliverables including: an inception report, the draft and final MTE plan, draft and final MTE reports, an executive summary of the final MTE report, presentations for specific audiences (including USAID), primary data sets, data collection instruments (English and all translations), descriptions of data analyses and results, lists of sites visited with types and numbers of informants at each, and any other required deliverables. This section must specify that the USAID evaluation policy requires that the report describe the strengths and limitations of the evaluation methods and how and to what degree these factors influenced the process and findings of the evaluation. The SOW should also specify that the report must clearly separate in different sections the evidence (i.e., raw data) collected by the evaluation team, the conclusions and recommendations that are based on the presented evidence. It should also state that sources of all evidence must be identified; conclusions must be based only on evidence

presented in the report; and recommendations must directly correspond to the conclusions. Finally, this section should articulate the timeline for completing the deliverables. The final report must adhere to the requirements in the USAID Evaluation Policy. Once approved by the AOR, the firm must submit the report, supporting documents, and related data sets to the partner in time for submission to BHA and the DEC within 30 days.

- *Pertinent Permissions, Approvals, Insurance, and Other Required Permits.* This section should indicate that the MTE team will be responsible for obtaining all necessary permissions, approvals, insurance, and other required permits and for adhering to national and local formalities. These include required permits related to data collection from human subjects, including necessary internal review board approvals and health and accident insurance for MTE team members.
- *Time Frame*. This section should identify the data collection period for the MTE, as agreed by the partner and BHA, and provide an illustrative time allocation for the various stages of the evaluation, subject to approval by the partner and USAID. This will help USAID, the partner, and MTE team estimate the financial resources needed for the evaluation. The time allocation should consider the number of team members and the accessibility of intervention sites.
- *Logistics*. Generally speaking, the research/evaluation partner will be responsible for the evaluation logistics. To avoid compromising activity implementation during the MTE and to maintain a separation between the MTE team and the implementers, activity vehicles and other vehicles branded to identify them with the partner or any of the implementing partners should not be used by the MTE team while they are in the activity area.

MTE Team Composition, Qualifications, and Roles

- *MTE Team Composition*. The MTE team will typically consist of a team leader who is an evaluation specialist, plus three to five technical specialists who bring expertise and practical experience in one or more of the activity's technical sectors and interventions in addition to strong qualitative research skills and experience. The team may include a data analyst experienced in analyzing and relating data across various technical sectors.
- *MTE Team Members' Qualifications.* The SOW should identify the following minimum requirements for MTE team members' qualifications:
    - Every team member's resume must show substantial application of qualitative research skills in developing countries.
    - The team leader must have significant formal education in a field relevant to evaluation (e.g., program evaluation, statistics, economics, agricultural economics, anthropology, applied research, organizational development, sociology, or organizational change) at a post-graduate or an evaluation professional continuing-education level.
    - The team leader must have extensive experience in evaluation using mixed methods of investigation (qualitative and quantitative) in developing countries. Knowledge of the conceptual framework of food security and experience evaluating food security programming is highly desirable.
    - Each technical specialist should have a post-graduate degree in a field related to at least one of the technical sectors of the activity, plus extensive practical experience in developing countries with interventions similar to those implemented by the activity.
    - At least one member of the team must have substantial demonstrated experience in gender integration.
    - The MTE team should comprise technical expertise from all activity sectors and activity management.
    - No member of the MTE team should have had any prior input to the activity's design or implementation.

- Having conflict sensitivity expertise on the MTE team is also desirable, particularly when an MTE will be undertaken in a context that is prone to violent conflict.
- Knowledge and experience with reviewing TOC and Logframe is highly desired.

- *MTE Team Members' Roles.* The SOW should describe each team member's role in the evaluation, as follows:
  - Team Leader: Organize and lead the overall evaluation; ensure a thorough review and analysis of activity monitoring data and other available secondary data by the appropriate team members; lead the selection of a purposively selected sample of implementation sites and Outputs for primary data collection; ensure an MTE plan that includes adequate triangulation and validation of evidence collected in all sectors; lead the collection and analyses of primary and secondary data to evaluate the activity's M&E processes and the integration of activity sectors and interventions; ensure that final report presentation is logical and presented in a way that clearly separates the evidence collected, conclusions, and recommendations in different sections of the report, and conclusions and recommendations are based only on the evidence presented in the report; lead writer who ensures the evaluation report is written clearly and concisely, organized and has a uniform voice; interact, on the part of the MTE team, with the partner and USAID; and serve as a technical specialist for specified sector (optional).

  - Technical Specialists: Lead the collection and analyses of primary and secondary technical data related to his/her field(s) of expertise, document findings, and draw conclusions and form recommendations for the sector(s); evaluate the general aspects of the implementation of all interventions related to his/her sector(s). While the team leader will likely be tasked as the primary investigator for the activity management overall, a technical specialist must consider management aspects of the implementation of interventions in his/her technical sector and the interaction between his/her technical sector and other activity sectors by examining: staff and material resources; communication, both internal and external; community involvement; participant targeting (especially overlap/consistency with other sectors); management of food and non-food commodities; transfers of entitlements (food, non-food, cash); branding; partnerships and linkages; consortium management; routine monitoring and data quality assurance for all interventions; exit/sustainability strategies; gender integration; environmental protection; and draft the report sections assigned by the team leader in the specified format.

Implementing Partner Roles & Responsibilities. This section should describe how the partner will or will not support the MTE team during the evaluation process.

- *Provision of Secondary Data.* This section of the SOW should list the resources that the partner will supply the team and the date when each resource will be available to the evaluators. To enable adequate time for secondary data analyses, the partner's staff should assemble maps, documents, databases, and other resources for the evaluators' use and deliver them to the evaluators at least two (2) months before the start of primary data collection, and the MTE team members should accomplish the review before arriving at the activity site. The contract should include at least two (2)–four (4) weeks of paid time during this period for each team member to review the secondary data. BHA recommends that partners collect and archive these materials throughout the life of the activity rather than waiting until the evaluation time approaches. This should be part of the M&E Plan.

- *Other Resources:* An illustrative list of resources that would be useful and should be available to the evaluators includes:
    - Lists of intervention sites, identifying the type(s) of interventions at each location, with start dates of implementation, numbers of direct participants and indirect beneficiaries, quantities of commodities distributed, etc., for each type. (This list should include locations of all community assets developed or rehabilitated using activity resources, including those still in process of development or rehabilitation.);
    - Food, voucher, cash, and non-food item distribution reports that include location of distribution; type of distribution; and planned and actual quantities, ration sizes, and timing of distributions;
    - Locations of all warehouses and identification of the managing organization, with uses and storage capacities;
    - Activity commodity management tools and reports;
    - Maps showing the activity area with administrative boundaries, roads, markets, food distribution points, intervention sites, partner offices, lodging, livelihood or ecologic zones, etc.;
    - Approved activity proposal narrative and relevant attachments with documentation of approved modifications;
    - Inception report;
    - Baseline study report and reports from all research conducted for the activity's benefit (e.g., formative research, barrier analyses, gender analyses, and market analyses);
    - A current organogram of activity staff (with names and phone numbers for incumbents and notation of vacancies) showing partner organization and supervision/management lines;
    - Intervention implementation protocols and guidelines and identification of activity staff who use each;
    - Descriptions, dates, and numbers of participants of capacity building activities for partner staff and activity participants (individuals, groups, and communities);
    - Complete M&E Plan, including monitoring tools, manuals, and reports;
    - Examples and lists of recipients of all types of M&E reports;
    - Activity monitoring databases;
    - Descriptions of the nature (e.g., format, location) and contents (e.g., type of data, period of collection) of the various data sets that the partner will provide for the evaluation;
    - Exit strategy and sustainability plan;
    - IEE, EMMP, and all related reports;
    - All ARRs, quarterly reports, and PREPs.

- *Logistical and Administrative Advice and Support*. This section should clarify what logistical administrative support the partner will provide and what the MTE team is expected to manage. For example, the partner may:
    - Arrange meetings between the evaluation team and USAID, at a minimum at the beginning and end of the evaluation process;
    - Provide contact details for key partners' staff;
    - Provide administrative support: communication, photocopying, printing, etc.;
    - Advise about local protocols and permissions to gain entry to operational areas;
    - Provide advice and support related to travel (international travel, travel routes, security conditions, local vehicles, and drivers for hire);
    - Identify local firms with potential to provide technical expertise, including translation, to the MTE team.

Ethical Guidelines. The SOW must clearly state that every member of the evaluation team must adhere to ethical guidelines as outlined in the *American Evaluation Association's Guiding Principles for Evaluators*. A summary of these guidelines is provided below.

- Systematic inquiry: Evaluators conduct systematic, data-based inquiries.
- Competence: The evaluation team possesses the education, abilities, skills, and experience appropriate to undertake the tasks proposed in the evaluation. Evaluators practice within the limits of their professional training and competence, and decline to conduct evaluations that fall substantially outside those limits. The evaluation team collectively demonstrates cultural competence.
- Integrity/honesty: Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.
- Respect for people: Evaluators respect the security, dignity, and self-worth of respondents, activity participants, clients, and other evaluation stakeholders. Evaluators regard informed consent for participation in evaluation and inform participants and clients about the scope and limits of confidentiality.
- Responsibilities for general and public welfare: Evaluators articulate and take into account the diversity of general and public interests and values that may be related to the evaluation.

Intellectual Property. The SOW must clearly indicate USAID's, the research/evaluation partner's, and the implementing partner's rights to intellectual property produced under the MTE. Unless otherwise provided in the award's provisions, the research/evaluation partner may retain the rights, title, and interest to data that are first produced. In addition, the SOW should state the following: "USAID reserves a royalty-free, worldwide, nonexclusive, and irrevocable right to use, disclose, reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, in any manner and for any purpose, and to have or permit others to do so."

Table A1. Illustrative Midterm Evaluation Question Matrix

Areas of focus	Aspects to consider	Illustrative methods of investigation/evaluation
1. How well have the activity's interventions met planned schedules, participant numbers, and Outputs? What factors promoted or inhibited adherence to schedules? How were problems and challenges managed?		
Adherence to planned schedules	• Start dates and rates of expansion of coverage for each intervention • Numbers and timeliness of planned participants and Outputs, e.g.: ○ Formative research, barrier analysis, gender analysis ○ Various types of direct trainings ○ Indirect training through trainees, e.g., farmer to farmer or cascade training ○ Formation of or connections with community groups ○ Construction or rehabilitation of assets ○ Development and progress of community action plans ○ Distribution of cash and goods	• Use secondary data from routine monitoring, ARRs, and other reports to compare planned and actual start dates, numbers of Outputs, and other targets, noting differences in achievements according to location, implementing partner, or sector. • Compare across locations, participant groups, activity administrative units, etc. to identify factors associated with differing degrees of achievement. • Interview members of activity staff at various levels about factors that delayed or interrupted interventions and Outputs, and how problems were identified and managed. • Ask groups and individuals from different stakeholder groups at locations of greater and lesser achievements about factors they believed inhibited or promoted efficiency and efforts have been made to overcome barriers.
2. What are the strengths of and challenges to the overall activity design, implementation, management, communication, and collaboration so far? What factors appear to promote or challenge the activity operations or effective collaboration and cooperation among the various stakeholders?		
Activity management	• Strengths and weaknesses of the activity work plan and schedule • Evidence that management has explored and implemented new and/or innovative ideas and approaches • Changes and challenges in the operating context and how management responded	• Review the activity work plan and schedule to assess how completely and clearly they define the work needed to meet objectives, when, and by whom. Is the schedule feasible? • Examine the roles of the different implementing partners and how the plan promotes good collaboration among them and leverages partners' relative advantages. • Interview members of management about Outcomes of work plan reviews and how they handled changes and challenges that presented. • Ask implementing staff in different roles how feedback and ideas are solicited and shared within and among partners, especially among field, country office, and headquarters.

Staffing	• Adequacy of numbers and capacities (knowledge, experience) from beginning of the activity until present • Strengths and weaknesses of supervision and support to ensure accountability, performance, and confidence among implementing staff • Adequacy or inadequacy of resources (tools, work space, transportation, communication, information, work aids) to support interventions' efficient performance at all times from start to current time • Gender sensitivity and balance at various levels of staff • Conflict sensitivity	• Review the characteristics and capacities of staff at all levels in all sectors, and assess their confidence and capacities to perform assigned tasks. • Review training and supervision schedules for monitoring and supporting implementing staff, including an assessment of the numbers of people and sites per supervisor. • Interview a sample of field staff and supervisors in different sectors and interventions and at different levels about: o Factors that affect their performance and motivation o Sources of satisfaction and dissatisfaction with resources, training, supervision, and support from technical experts o Their roles in decisions about intervention design and implementation o Training they received
Community engagement and participation	• Strengths and weaknesses of the ways the activity has included community members, including vulnerable or marginalized members, in decisions about intervention choices, design, implementation, and monitoring • Community members' perceptions about their participation in the activity and the degree and nature of their engagement with activity staff • Use of incentives of different types • Safeguards against exploitation and discrimination • Application of "do no harm" principles	• Interview groups and individuals from different stakeholder groups about: o Who has been involved in the activity and how? o Who else would have liked to be involved, and how? o Satisfaction and dissatisfaction with the way activity staff and community members engaged o Reports of or potential for exploitation or discrimination by implementing staff or with activity resources • Interview volunteers who contribute significant amounts of unpaid time to activity-supported interventions to learn what motivates their participation and performance.
Cultural acceptability	• Implementation methods: type, timing, style • Interactions between implementing staff and community members • Messages, and methods and timing of communication • Outputs	• Interview members of target communities, government counterpart agencies, and field staff to assess perceptions and attitudes about the choice, implementation, and Outputs of interventions: o How well do the interventions address perceived needs? o What aspects do they like or dislike? o How would they prefer things to be done?
Communications	• Quality and timeliness of communications of vision, objectives, plans, implementation guidelines, and other activity information	• Interview members of implementing partners, communities, government counterparts, and other stakeholder groups to assess:

	among activity staff, partners, government counterparts, and communities • Knowledge in various stakeholder groups about the activity • Strengths and weaknesses of the ways the activity encourages and handles feedback from community members, staff, and partners	○ Knowledge of objectives, interventions and implementation, intervention duration, eligibility, Outputs, and entitlement transfers ○ How and when they learned about activity objectives and interventions ○ Frequency and content of communications with other types of stakeholders ○ Satisfaction and dissatisfaction with ways feedback is received and responded to
Partnerships and linkages	• Consider collaboration and links with: ○ Other USAID activities ○ Government activities ○ Community based organizations ○ Other complementary activities in the activity area • Strengths and weaknesses of coordination within the activity and between the activity and other activities and agencies • Factors that make partnerships more or less beneficial to activity implementation	• Interview implementing staff, government counterparts, members of community organizations, and staff of linked or collaborating activities about: ○ The nature and sources of satisfaction and dissatisfaction with their collaboration and ways that it could be improved ○ How they feel their cooperation benefits the implementation and results on both sides ○ Other activities, agencies, and groups that are doing similar or complementary work to which the activity is not linked • Review samples of activity's memorandums of understanding with collaborators
Financial management	• Financial accountability • Sufficiency of finances to ensure good activity implementation • Flexibility of the budget to respond to changing conditions	• Review financial records • Interview managers about: ○ The adequacy of finances and effects of financial constraints on activity implementation ○ Perceived limits of financial flexibility to respond to change
Branding	• Compliance with USAID policy • Knowledge and attitudes toward donor and implementers within target communities	• Assess how well planned and actual actions do or do not comply with USAID branding requirements. • Interview participants and community leaders about their knowledge of and attitudes toward USAID and implementing partner agencies.
M&E	• Completeness and clarity of the documented M&E Plan. • Ways the M&E system: collects data useful to monitor the quality and Outputs of processes; solicits and reports opinions, ideas, and concerns from field staff; provides constructive feedback to implementing staff to inform, assist, and ensure accountability and	• Critically review the M&E Plan and systems: staffing, processes, and Outputs. • Interview staff in various roles in the collection, analysis, and reporting of routine monitoring about their activities and roles, to determine their understanding and confidence in the data collected, and challenges they face getting or using the data. • Interview recipients of reports and other Outputs about how they use the information they receive, which information is most useful, the timeliness of the information, and any other information they would like to have.

	motivate good performance; ensures accurate reporting to USAID; supports timely problem solving and decision making for all stakeholders; ensures data quality: validity, reliability, timeliness, integrity, and precision; has been used to adjust implementation; is or is not supported by the institutional structures; monitors environmental impact; monitors gender equity; and monitors context; monitors unintentional results (positive and negative). • Strengths and weaknesses of data collection methods • Design; management; and roles in monitoring, analysis, and report generation of databases. • Challenges the M&E team faces	• Interview key decision makers about the timeliness and usefulness of the data from the M&E system.
Environmental safeguards and compliance	• Adequacy of the EMMP • Adherence to the details of the EMMP through specific environmental monitoring systems • Incorporation of the EMMP into the IPTT and monitoring processes • Recognition or avoidance of unforeseen environmental damage and climate stressors	• Examine how well planned and actual actions and Outputs do or do not comply with the activity's EMMP. • Interview technical experts, implementing staff, and other key informants about activity interventions' apparent or potential threats to the environment and identify those not addressed by the EMMP and how well the activity implementation has addressed these threats.
Commodity management	• Division of roles and cooperation among activity staff, external partners, and community members • Common causes of delivery delays and commodity losses • Adequacy of mechanisms to safeguard against loss or abuse • Adequacy of mechanisms to ensure adherence to "do no harm" principles • Quality and cleanliness of storage facilities • Completeness and clarity of commodity records and reports	• Review processes and records related to tracking, management, and delivery of commodities to points of use or distribution: identify challenges and measures taken to address them. • Inspect warehouses and storerooms of various sizes and managed by different entities. • Interview people managing the resources about what is and is not working well, why, and how problems have been addressed.

Exit and sustainability strategies	• Comprehensiveness of the exit and sustainability strategies • Factors that threaten the continuation of targeted practices and services and the maintenance of new infrastructure • Progress in implementing the strategy • Ways the activity is strengthening or establishing links between communities and private or public financial or technical resources	• Critically review the exit and sustainability strategies and progress in its implementation in light of the findings related to the challenges to practices promoted by and threats to infrastructure developed by the activity. • Interview key informants and participants about threats and promoters of targeted practices and infrastructure.
3. In each technical sector, what are the strengths of and challenges to the efficiency of interventions' implementation and their acceptance in the target communities? How well do implementation processes adhere to underlying principles and activity protocols? What factors in the implementation and context are associated with greater or lesser efficiency in producing Outputs of higher or lower quality? Which interventions and implementation processes are more or less acceptable to members of the target communities and why?		
• Behavior change communication • Direct and indirect training • Health and nutrition • Water, sanitation, and hygiene • Agricultural production • Income generation • Savings and loans • Natural resource management and environmental protection • Women's empowerment and promotion of gender equity • Prevention of gender-based violence	• Application of findings from formative research and gender analyses to implementation • Technical quality of activity Inputs and Outputs • Strengths and weaknesses of how the various interventions engage target groups and protect against unintentional harm • Selection of direct beneficiaries; coverage of target groups • Perceptions of quality, appropriateness, and use of distributed goods and promoted services • Composition, activities, and governance of groups created or promoted by the activity • Networks and connections facilitated by the activity • Collaboration with and support to relevant government service providers • Cultural acceptability and relevance of intervention methods and messages	• Review formative research and evaluate how well implementation has applied the findings. • Observe interventions (training sessions, distributions, construction, community meetings, FSS or care group sessions…) and talk with implementing staff and direct participants about: o What interventions are more and less effective o What and how could interventions be improved o Which interventions are more or less interesting or useful o Who benefits; who should benefit; how are participants selected o Opportunity costs of participation in interventions o Knowledge and understanding of key activity messages • Talk with non-participants from the same communities about: o Which interventions are more or less interesting or seem more or less useful o Who benefits; who should benefit; how are participants selected o Perceptions about the benefits they could gain with participation in interventions • Compare and contrast men's and women's participation and perceptions. • Review messages on the same topic transmitted through different pathways for consistency and clarity. • Compare the understanding of the key messages of trainers and direct and indirect trainees. • Inspect the technical quality of community and household infrastructure and natural resources to which the activity contributed.

• Early warning systems and disaster risk reduction • Community governance • Food/cash for work	• Consistency of content and recipients' understanding of similar messages received via different pathways • Strengths and weaknesses of measures taken to ensure gender equity with regard to access to, participation in, and benefit from activity interventions • Strengths and weaknesses of linkages, coordination, and integration among the different sectoral and cross-cutting technical areas • Validity and comprehensiveness of assumptions in the activity's TOC that are critical to intervention implementation and Outputs	• Examine the composition of the various groups created or supported by the activity: Who in the community did or did not join. Why or why not? How is the gender balance? Are marginalized groups represented? • Talk with members of groups formed or promoted by the activity about: how the group was formed; level of satisfaction with group composition; challenges and successes working as a group; nature and adequacy of support from program; how and why the group chooses interventions; how members' other roles in the community affect their participation in the group. • Examine participant records to assess the proportion of households and communities that benefit from multiple sectors in different combinations. • Interview members of households benefiting from interventions in single and multiple program sectors about their participation; compare characteristics of those who benefit from one vs. multiple sectors. • Interview staff members about their interactions with staff working in other sectors, especially regarding site and participant selection and developing messages to participants. • Interview members of various types of groups initiated by the activity about, e.g., making decisions, managing joint resources, and sharing information and experiences.
4. What changes—expected and unexpected, positive and negative—do community members and other stakeholders associate with the activity's interventions? What factors appear to promote and deter the changes? How do the changes correspond to those hypothesized by the activity's TOC or LogFrame?		
Changes observed or reported	• The activity's TOC and Logframe • Intended and unintended change • Positive and negative change • Differential change among participants (individual, community) of one sector, participants of multiple sectors, and non-participants • Differential change among participants representing different population sub-groups • Perceived benefits of participation in interventions from multiple sectors vs. a single sector • Perceived trajectory of change and conditions that threaten or promote sustained change • Changes in conditions related to assumptions	• Interview community members (participants of one or more sectors and non-participants) and activity staff to gain perspectives about: changes they have made themselves, observed in others, or observe in the social, economic, or physical environment; factors that promoted the changes; barriers to changes intended by the activity; conditions that promote or threaten sustained change • Technically evaluate how strategically selected infrastructural Outputs affect or can affect livelihoods, well-being, or environmental conditions.

5. Based on the findings from Questions 1–4, how could the activity be modified to improve its acceptability to targeted communities or the efficiency and effectiveness? How should the activity's TOC or LogFrame be refined or modified?		
Based on findings from 1-4 above	• Observed and perceived strengths and weaknesses of the implementation so far • Factors in the design, implementation, and context that affect the efficiency or acceptability of the processes, Outputs, and Intermediate Outcomes • Targeted communities' and individuals' perceptions and priorities • Relative cost and feasibility and anticipated value of acting and benefiting within the life of the activity • Potential to advance the activity's ultimate objectives and Goal	• Use the results of inquiries to the questions above to form conclusions and recommend concrete actions to help improve activity performance and final results. • Prioritize the recommendations and identify the actor(s), the purpose for change, and anticipated benefits. All recommendations should be directly related to stated conclusions and based on evidence presented as findings.

ANNEX III. INTERIM/FINAL EVALUATION OVERVIEW

This annex provides a general overview of the interim/final evaluation and the endline study methodology in particular. It is intended to: (1) inform the development of a scope of work for the research/evaluation partner conducting the baseline study; and (2) to help BHA resilience, recovery, and preparedness activity implementing partners better understand the purpose, scope, and methods used for interim/final evaluation.

Note that this annex primarily refers to endline study methodology used for a pre-post final performance evaluation design. Some of the basic principles and considerations included here may be helpful to inform endline data collection for an experimental or quasi-experimental evaluation design.

Overview

BHA will manage an external research/evaluation partner to conduct the interim/final evaluation. Before commencing the quantitative endline survey, the research/evaluation partner submits to BHA a survey design protocol that includes:

- A sampling plan, including sample design, sampling frame, sample size calculation, level of statistical precision and power, and respondent selection procedures;
- Training and field manuals for supervisors, enumerators, and anthropometry;
- Plans for supervisor, enumerator, anthropometry, and data entry training;
- A questionnaire translation, back-translation, pretesting/piloting plan; and
- Data analysis and treatment plan, including estimation procedures (sample weighting and other adjustments) and indicator tabulation and subgroup analysis.

Implementing partners will brief the endline study research/evaluation partner on the cultural, political, and social contexts in which it will be collecting data to help it develop a survey design that will do no harm. The research/evaluation partner should also demonstrate how "do no harm" principles apply. (The research/evaluation partner should also summarize in the interim/final evaluation report the steps it took to maintain conflict sensitivity of its approaches, in line with what was completed during the baseline study.)

The research/evaluation partner must submit to BHA the survey design for approval to the AOR. The survey may commence only after BHA has approved the final survey design.

After the study is complete, the interim/final evaluation research/evaluation partner will submit a draft report to the awardee and BHA for comment. After responding to the comments, the research/evaluation partner will submit the final report for approval. Once approved by the AOR, the research/evaluation partner must submit the report, supporting documents, and all related data sets to the awardee in time for submission to BHA and the DEC and DDL within 30 days.

Final Evaluation Objectives and Questions

The ultimate objective of the interim/final evaluation is to inform BHA's decision to award a cost extension. To that end, the interim/final evaluation may include the following illustrative questions, methods, and approaches:

Question 1. *To what extent has the activity met its defined goal, purposes and outcomes?* To achieve this objective, the evaluation will use both quantitative and qualitative methods while taking into account (a) activity's performance on indicators against targets jointly set by the implementing partner and BHA; (b) factors that promoted or inhibited the achievement of the activity's results; (c) plausibility of pathways and the determinants of achieving the key outcomes; (d) targeting strategies and their contribution to achieving activity goal, purposes, sub purposes, and intermediate outcomes (especially with regard to gender and reaching the most vulnerable); and (e) the appropriateness and effectiveness of interventions on the poorest individuals.

Question 2. *Based on the evidence, which outcomes are likely to be sustained and what systems, capacities, linkages, resources, motivation and processes were developed to sustain them?* To achieve this objective, the evaluation will examine the functionality and performance of systems and processes established independently by the activity, as well as in collaboration with the private sector, government departments, and research organizations to achieve food security outcomes and sustainability. Using primarily qualitative methods the team will explore (a) the quality of the processes, systems, governance, and institutional arrangements developed and/or strengthened to sustain the necessary and critical services; (b) communities' perceptions about the quality, frequency, effectiveness, and sustainability of the services provided by the activity; (c) the likelihood that service providers will continue providing services after the activity ends; (d) the motivation of the community and participants to demand and pay (or invest time) for the services; (e) whether the necessary resources and capacity strengthening support will exist to sustain service providers' needs; (f) the extent to which the activity leveraged other USG and non-USG investments to achieve sustained outcomes as identified in the theories of change; and (g) evidence of enhanced linkages with other service providers.

Question 3. *In each technical sector, what are the strengths of, and challenges to, the effectiveness of implementing the interventions, and the acceptance of interventions in the target communities?* To achieve this objective, the evaluation will assess the effectiveness and relevance of the technical interventions to achieve food security outcomes, and discuss those findings in relation to the activity's theory of change. The team will support its determination using both quantitative and qualitative methodologies when discussing the following: (a) factors in the implementation process and context associated with greater or lesser efficiency and effectiveness in producing outputs of higher or lower quality; and (b) the interventions and implementation processes deemed more/less acceptable to members of the target communities.

Question 4. *How has the management of the activity contributed to/impeded the achievement of results?* Areas of exploration may include financial management, staffing (both at HQ and in the field), communication within a consortium, and partnerships with key actors. To achieve this objective, the evaluation will assess the structure, capacity, system, and functionality of the activity management. The evaluation team should assess the functionality of the feedback loop to get input from staff and partners, how the management system has adapted to the need of an effective and high quality implementation of interventions. This could include how the activity and the prime organization responsible for hiring and management of staff; staff capacity development; investments in M&E systems; and data utilization for adaptive management.

Question 5. *What key lessons learned and best practices should inform future activities in the BHA programming?* To achieve this objective, the evaluation will identify best practices, strengths, and challenges in the activity design (including the theory of change), that support the activity's achievements,

as well as approaches that should be considered in designing future food and nutrition security activities, and strengthening household and communities' resilience capacities. The evaluation team will use both quantitative and qualitative methods is assessing (a) the unintended positive and/or negative consequences of the activity's interventions; (b) ways to minimize potential unintended negative consequences and systematically capture positive consequences.

Evaluation Methodology

All final evaluations will include a preliminary desk review then mixed-methods data collection and analysis.

Desk Review. The implementing partner should provide the research/evaluation team with following documents for the desk review: activity proposal; theory of change and M&E Plan; PREPs; Annual Results Reports; Indicator Performance Tracking Tables; Midterm Evaluation report and Post-Midterm Utilization and Action Plan; baseline study report; Implementing partner formative research studies and/or any other qualitative studies that were carried out during implementation; monitoring data and and success stories or other reports generated by the implementing partner; Initial Environmental Examination Report and Environmental Status Report; any other national level reports such as Demographic and Health Survey, Living Standards Measurement Study, agricultural censuses or data that may be relevant.

Quantitative Endline Survey. The endline household survey will collect data on the same indicators that were collected at the baseline study. The timing of the endline survey typically aligns with the baseline survey to allow comparability of results. The evaluation team will typically use the same sampling frame, data collection instruments, level of statistical precision (95 percent confidence intervals), and statistical power (80 percent) as the baseline study; however, the quantitative endline survey design may not necessarily be identical to the baseline survey. For example, if the activity reduced or expanded its target areas, the sampling frame may need to be adjusted.

Qualitative Inquiry. Qualitative methods will be used to collect information to interpret quantitative results and to answer evaluation questions that are qualitative in nature. The evaluation team will design the overall qualitative study approach and should consider a variety of primary data collection methods, such as semi-structured in-depth interviews, group discussions, focus group discussions, key informant interviews, direct observations, and case studies. As with the quantitative household survey, qualitative sampling should include both individuals who directly participated in the activity and those not specifically targeted with any intervention. The purpose of interviewing non-participants or indirect-participants is to understand whether non-participants acquired knowledge, techniques, and behaviors from the activity promoted interventions, and whether the activity negatively affected community members who did not directly participate in the activity's interventions. In addition, the evaluation/research team should interview USAID personnel, project staff, knowledgeable people from the community, local government staff, community leaders, host government officials, and staff from other agencies and individuals as appropriate.

Data Analysis

For pre-post performance evaluations, the evaluation team will statistically compare (using tests of difference) the endline data to that of the baseline to detect changes (if any) for all key indicators. The evaluation team will conduct descriptive and inferential analyses to describe the results, detect differences, as well as various econometric analyses to identify the determinants of key outcomes and the magnitude and direction of changes. In advance of fieldwork, the evaluation team should develop a data analysis plan and submit this to BHA for approval. When analyzing the data, however, the evaluation team will not limit itself to the data analysis plan; rather, the evaluation team should keep an

open and curious mind to look for correlations between variables. For impact evaluations, the data from the treatment/ program communities will be compared with the data from the control communities.

Final Evaluation Timeline

From the planning phase to the final report, a performance evaluation typically takes up to 15 months. This time period includes activities such as finalizing the scope of work, data collection instructions, field-based data collection, data cleaning, data analysis, triangulation and merging of qualitative and quantitative findings, and drafting/reviewing the report. The exact timing of a final evaluation will depend on a number of factors, including BHA's plan to continuously invest in the country, data needs for activity design, performance of the activity as determined by the annual results report, field visits, monitoring reports, and midterm evaluation report. Preparation for a final performance evaluation should begin at least one year before the results are needed. An illustrative timeline of deliverables is as follows:

- Drafting and approval of SOW: 12–16 weeks
- Procurement of evaluators: 12–16 weeks
- Desk review and work plan preparation: 3–5 weeks
- Data collection: 6–8 weeks, excluding travel time
- Data analysis and drafting of report: 6–8 weeks
- Feedback from USAID and finalization of final evaluation report: 4–8 weeks

Evaluation Team Composition

A final evaluation team consists of a team leader, subject matter specialists, a survey specialist, and a survey team. All members of a final evaluation team will be external to the implementing partner and its consortium members. No member of the final evaluation team may have had any responsibility in the design or implementation of the activity under evaluation. The team leader must be an evaluator with demonstrated expertise and experience in leading at least five evaluations of multisectoral nutrition and food security programs. The team as a whole will comprise of subject matter experts and have expertise in all of the activity's technical sector intervention areas and cross-cutting technical areas.

To avoid disruption of activity implementation and avoid potential bias that could affect the evaluation results, the final evaluation team must not use activity staff as translators, enumerators, or supervisors to carry out any data collection. The evaluation/research team may, however, interview activity staff members, the responsible BHA Officer, the AOR, and any other USAID or partner staff member with a direct stake in the activity as key informants. Partner and USAID staff may also review and provide comments on data collection tools and instruments before they are finalized.

ANNEX IV. MANAGEMENT INFORMATION SYSTEM (MIS) KEY PRINCIPLES AND BEST PRACTICES

The purpose of this annex is to present BHA recommended key principles and best practices for designing and operationalizing a Management Information System for BHA funded activities.

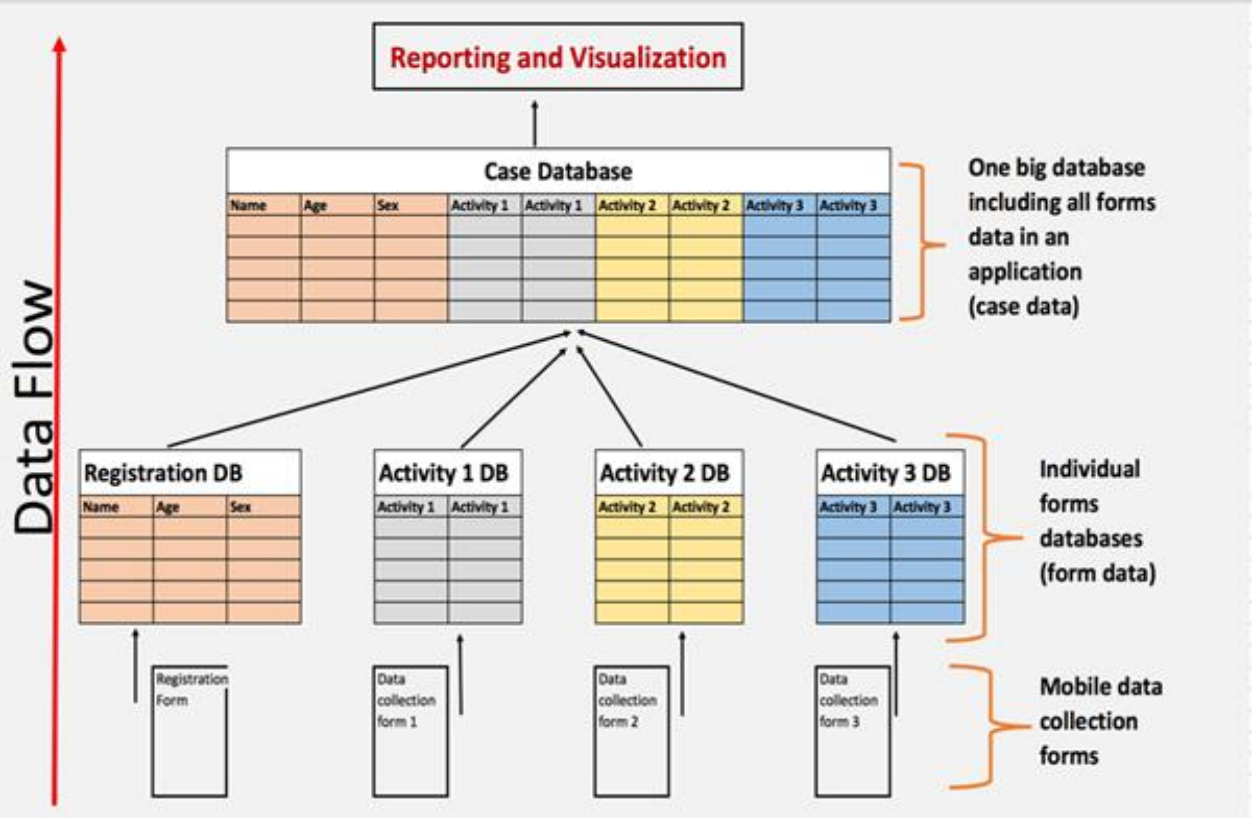
Key Concepts and Design Considerations

Management Information System: BHA defines MIS as a set of data collection, storage, and reporting software tools for reporting and decision making. Often development organizations use the terms MIS, monitoring database, monitoring system, M&E system interchangeably.

Mobile based data collection: BHA recommends mobile based electronic data collection, as much as possible. Partners should use mobile devices for collecting monitoring data regularly synced to the cloud for safe storage and further processing. There are several ready-to-use software services available that enable field data collection in offline mode.

Case Management Approach: BHA recommends that partners use a case management approach while designing an MIS. The case management approach has several advantages for multi sectoral interventions like the elimination of double counting, reduced data collection time, and significant reduction in time for data analysis, etc. The basic principle is to collect necessary identification and demographic data during participant registration and then collect activity specific data during routine monitoring. For example, a patient provides several details about herself on the first visit to a new doctor, e.g., ID, name, age, sex, address, etc. For all future visits, the doctor's office only asks for her ID before recording the current visit related data: date, time, and reason for the visit, etc. Similarly, public health interventions use case management for repeat house visits; emergency responses use it for multi-sectoral distributions to the same or partially overlapping participants.

Relational Database: A relational database is used as a technical backend for a case management system. The user interface (front end) of an MIS has several interrelated data collection forms which are stored as a web of interrelated data tables. The relational aspect of case management approach enables the system to create unique IDs, identify duplicate records, and eliminate double counting of participants. The following is a simplistic illustration of how multiple data collection forms save data in separate data tables in a database which then can be exported in one large dataset for further analysis and reporting.



Best Practices

Development awards are complex interventions that require a lot of coordination and logistics to deliver the right services to the right people. Creating an MIS for such interventions can be daunting especially knowing that several indicators change and/or archive over the activity period that require continuous updates in the MIS.

Stakeholders: Often designing a technology enabled product like a monitoring system is considered an 'IT task'. On the contrary, BHA thinks the non-IT program personnel should lead the MIS 'design' process and prepare a blueprint which then can be 'developed' by the IT teams. The program teams better understand the program logic, field realities, and data collection requirements, whereas the IT teams are better at weighing in on the type of software, mobile devices, database, and data security protocols. It is crucial for both parties to understand their roles in this process.

Existing monitoring system: BHA discourages creating a new MIS for each new activity. Partners are encouraged to look into previous MISs used by themselves or other organizations for prior awards and see how they can adapt them for new awards.

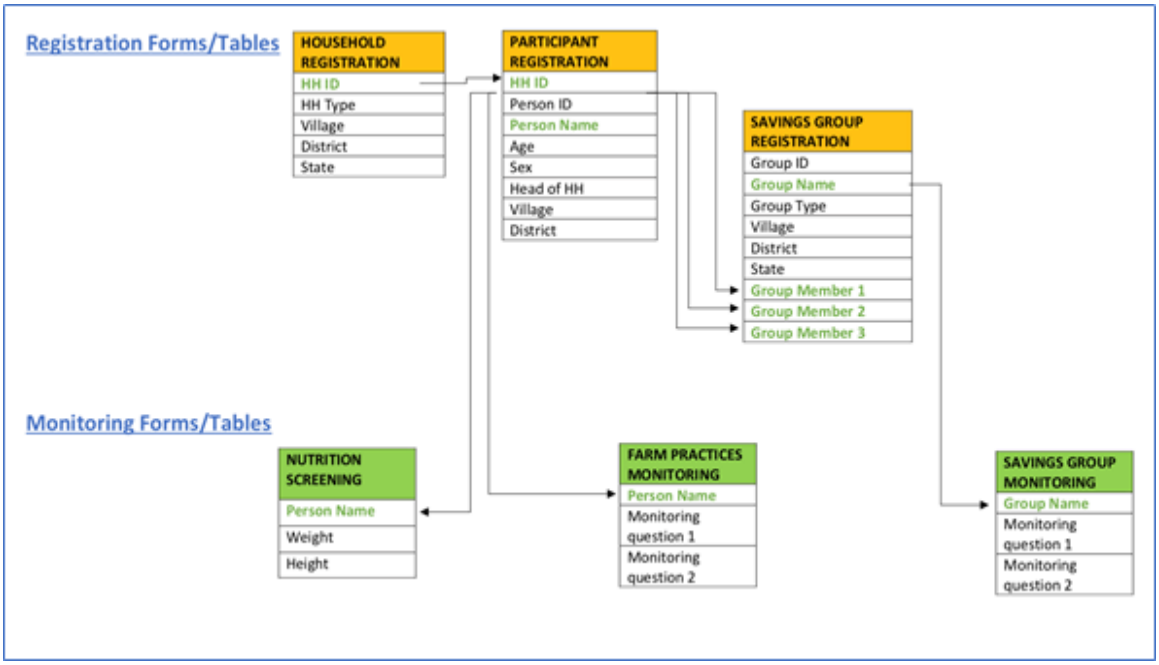
Common system: One of the key objectives of a relational MIS is to eliminate duplicate records or double counting. A time-tested good practice to achieve this objective is to develop common monitoring systems. For activities with a consortium of partners, the prime and sub-awardees should develop a common system instead of each partner using their own. Similarly, an activity should have a central MIS for all sectoral activities, instead of separate systems for each sector.

Existing corporate software licenses: Before purchasing software licenses just for a single activity MIS, the partners should communicate with their home office IT staff to find if they already have an

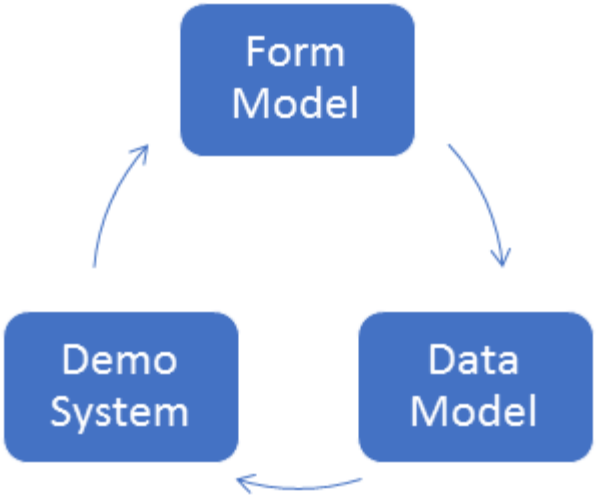
enterprise subscription that can be used. Often corporate licenses are cheaper than buying a few individual licenses.

Form Model: Once the program teams have all data collection forms ready, they should draw a form relationship model, illustrating a 'relationship network' between different data collection forms. This model is a precursor to a technical relational Data Model (discussed in the next section). Invite the field monitoring and IT staff to help design the form model. Do not try to finalize the Form Model in one design meeting. It will take a few iterations. You can reduce the number of design stakeholders with every successive iteration.

Data Model: A Data Model illustration is a more detailed and more technology oriented view of a Form Model where the relationships are visualized between the backend data tables. Please see a simplified example of a Data Table.



Test the MIS: Partners should field test the initial iterations of the MIS as early as possible to ensure users (e.g., field staff) are actually able to successfully use the MIS tools as they have been conceived of and designed.



Roles and responsibilities: Clear roles and responsibilities are crucial for the successful adoption of an MIS in any organization. Best practices include hiring a dedicated information manager or database officer for the activity. Modern MISs are complex and must be managed by specialized persons with degrees in computer science, data science, or both. This person should be seen as an MIS gatekeeper who processes, cleans, and exports required data for further analysis, and communicates to IT staff any changes in the database. Finally, let the IT team manage mobile devices and software troubleshooting.

Digital Development Principles

BHA recognizes the context for each activity is different, resulting in varying MISs from activity to activity. Nevertheless, BHA understands the importance of minimum digital standards and recommends following key principles from the USAID Digital Strategy and Principles for Digital Development.

Design with the user. Involve multiple user types and stakeholders in each phase of the MIS development to garner their feedback and ensure utility and applicability. For example, seek input from program/field staff, technical staff, host gov officials, implementing partners, and other stakeholders.

- Design tools that improve users' current processes, saving time, using fewer resources and improving quality.
- Develop context-appropriate tools informed by users' priorities and needs, considering the ecosystem and accepting that digital tools will not always be the best fit.

Understand the environment/context. Well-designed initiatives and digital tools consider the particular structures and needs that exist in each country, region and community. Dedicating time and resources to analyze the context where you implement, helps to ensure that selected technology tools will be relevant and sustainable and will not duplicate existing efforts. Context includes the culture, gender norms, political environment, economy, technology infrastructure and other factors that can affect an individual's ability to access and use a technology or to participate in an intervention. Context may change during the award timeframe, requiring the partner to regularly analyze the context to check their assumptions, monitor the context for changes throughout the life of the award, and adapt tools, processes, systems as needed.

- Engage with your target users and consult existing research to develop an understanding of the people, networks, cultures, politics, infrastructure and markets that make up your country context before designing your tool, process, or system.

- Coordinate with other implementing organizations, civil society and the government early on to learn from successful and unsuccessful initiatives/interventions in the implementation area, to avoid duplicating efforts and to integrate with existing technical systems more easily.
- Ensure that your tool, or process, system aligns with existing technological, legal and regulatory policies, best practices, and that you consider your organizational policies.

Data driven. Data use should drive what data to collect. No amount of data will lead to accelerated impact if it is not used to inform decision making. When a tool, process, or system is data driven, quality information is available to the right people when they need it, and they are using those data to take action. Data collected, entered, synthesized, processed and stored should be used for more than just outputs, such as published work or donor reporting. Examples of the types of data that can be collected to inform decision making include surveillance, research, operations, project management and data from secondary sources collected outside of the program.

- Collect data for outcome indicators, not just outputs.
- Use rigorous data collection methods. Consider and address potential biases and gaps in the data collected, perform data quality checks, and maintain strong documentation of collected data, including digital data sources.
- Use quality real-time or timely data to support rapid decision making, improve programming for users and inform strategy, if applicable.
- Present data in formats that are easy to interpret and act on, such as data visualizations.
- Create a data use culture by prioritizing capacity building and data use efforts across all stakeholder groups, including the groups whose data are being collected.
- Be holistic about data collection and analysis. Collect data from multiple sources and use a mix of data collection and analysis methods. Analyze your data collaboratively with stakeholders, if appropriate and applicable.

Reuse and improve. Instead of starting from scratch, activities that "reuse and improve" existing tools, processes, or systems will reduce the time needed for development and testing and reduce costs. Reuse means assessing what resources are currently available and using them as they are to meet program goals. Improve means modifying existing tools, products and resources to improve their overall quality, applicability and impact. Start by identifying relevant methods, standards, software platforms, technology tools and digital content that have already been tried and tested within and outside your organization. While an existing tool or system may not exactly fit all your needs for reuse, consider improving and building on it, rather than creating something entirely new. The result is a tool or system that is now better and more reusable by all because of your improvements.

- Identify the existing technology tools and systems (local/global and internal/external) within the food security community and within your geography and/or in your sector. Evaluate how these could be reused, modified or extended for use in your activity.
- Develop modular, interoperable approaches instead of those that stand alone. Interoperability will ensure that you can adopt and build on components from others and that others can adopt and build on your tool in the future.

Address Privacy and security: Partners must take measures to minimize collection and to protect confidential information and identities of individuals represented in data sets from unauthorized access and manipulation by third parties. Using data responsibly ultimately requires balancing three broad thematic areas:

- Data use — Data can be used to maximize the efficiency and effectiveness of programs and activities, with the goal of improving outcomes. They can help us target interventions to the

people or communities who will benefit the most. When shared with others, data can help us to build support and consensus by communicating challenges and possible solutions.

- Privacy and security — Data carry substantial risk, both for data subjects and for data stewards (e.g., implementers and donors). Partners are responsible for taking appropriate measures to minimize the risk to individuals based on data that is collected.
- Transparency and accountability — Partners are also responsible for sharing information with the people affected by our projects, host-country governments, and the U.S. taxpayers who fund our work.

These three areas of privacy and security are frequently in tension with one another. Understanding these tensions and working to balance them can help us work responsibly and highlight questions about risk and benefit surrounding our data. For example, a single-minded focus on data use might lead to over-collection of sensitive data, leading to potential privacy risks. On the other hand, if we prioritize privacy and security above all else, then we might delete data soon after collection. If done carelessly, this could compromise the records retention necessary for transparency and accountability. See Considerations for using Data Responsibly at USAID for more information about using data responsibly.

- Identify which data are collected and how data are acquired, used, stored and shared.
- Define data ownership and access before any data are collected or captured. Determine what local data protection laws and regulations need to be followed, who gets to decide what to do with the data, who is allowed to access or use the data and where data can (or must) be stored.
- Keep the best interests of end users and individuals whose data are collected at the forefront of your planning for upholding user privacy and ensuring data security and ethical implementation. This is especially important when implementers work with vulnerable or marginalized communities who may not have had a say in how their data have been collected, used or shared.
- Perform a risk-benefit analysis of the data being processed that identifies who benefits and who is at risk. This process may need to be repeated throughout the period of performance as new data are needed, new risks are identified or emerge, or new data-sharing partners are considered.
- Assess the risks of unauthorized access or leakage of any stored data. Consider the impact this data could have on the individuals if accessed or published maliciously and the risks if data were combined with other data sets.
- Understand that risks are highly contextualized, not just to countries but also to communities, populations and periods of time. If working with vulnerable or marginalized communities, what groups might have motivation to acquire your data, how capable are they, and are the information and access controls around the data sufficient?
- Minimize the collection of personal identifiable information. Consider how critical personal information is to the activity's success and what the consequences would be if those data are exposed to third parties — especially when partnering with users from vulnerable populations, such as minority groups, the disabled, and women and children. Include a risk assessment for collecting personal information.
- Be transparent with individuals whose data are collected by explaining how your initiative will use and protect their data.
- Obtain informed consent prior to data collection. It is crucial to ensure that participants understand why their data are being collected, how data are used and shared, and how the participants can access or change the data collected, and that they be given the option to refuse to participate. Participants should be informed of and fully understand the risks related to sharing their data. Consent forms should be written in the local language and easily understood by the individuals whose data are being collected.

ANNEX V. QUALITATIVE INQUIRY PLANNING SHEET (QUIPS) TEMPLATE

<table>
<tr><th colspan="2">QM/QS #: [Insert title of qualitative study or monitoring inquiry]</th></tr>
<tr><td colspan="2">PURPOSE and OBJECTIVES:</td></tr>
<tr><td colspan="2">Research/inquiry question(s):</td></tr>
<tr><td colspan="2">Data type (Process/Output/Outcome/Impact/Context/Cross-cutting theme):</td></tr>
<tr><td colspan="2">DESIGN and METHODOLOGY:
Data source(s) and research methods:

Sampling strategy & selection criteria:

Data collection tools, and implementation plan:

Frequency and timing:

Training requirements:

Data recording, data management, and quality assurance:</td></tr>
<tr><td colspan="2">DATA ANALYSIS PLAN:
Disaggregated by (e.g., gender, age, degree of poverty, family composition):</td></tr>
<tr><td>DELIVERABLES:</td><td>UTILIZATION/ APPLICATION:</td></tr>
<tr><td colspan="2">LIMITATIONS AND RISKS:</td></tr>
<tr><td colspan="2">ETHICAL REVIEW STATUS/ INFORMED CONSENT:</td></tr>
<tr><td colspan="2">IMPLEMENTATION TIMELINE:</td></tr>
</table>