



# Evaluation Planning Tool for USAID Nutrition Programs

USAID defines evaluation as "...the systematic collection and analysis of information about the characteristics and outcomes of strategies, projects, and activities as a basis for judgments to improve effectiveness, and timed to inform decisions about current and future programming. Evaluation is distinct from assessment or an informal review of projects" (USAID 2016, p. 3). Evaluation seeks to understand what a program accomplished.

USAID uses two broad categories to refer to program evaluation: performance (i.e., what was done) and impact (i.e., attribution). (See Box I). This guide presents considerations for nutrition program evaluations, including a decision tree and description of options for different types of evaluations, differentiating among them by methodology, questions that can be answered, and resource considerations to aid decision making. Common considerations for five key evaluation components are described:

- Program design and evaluation: What do you want to know? Who can find out?
- Identifying evaluation questions and appropriate methodologies
- Identifying appropriate indicators
- Analyzing data and producing the report
- Disseminating and using information

## Box I. Main Categories of Program Evaluation

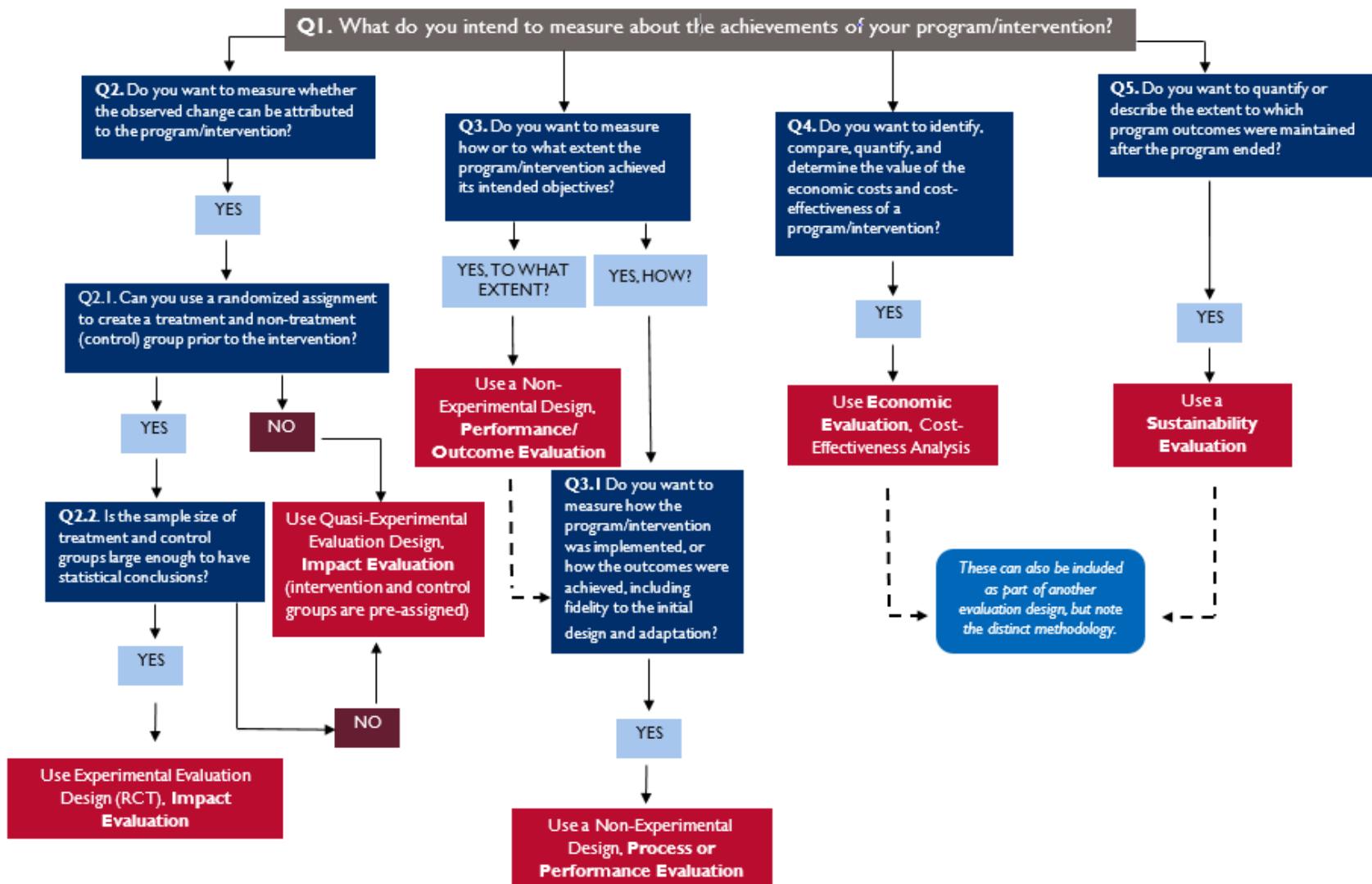
**Impact evaluations** measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.

**Performance evaluations** encompass a broad range of evaluation methods. They often incorporate before-after comparisons, but generally lack a rigorously defined counterfactual. Performance evaluations may address descriptive, normative, and/or cause-and-effect questions: what a particular project or program has achieved (at any point during or after implementation); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to design, management, and operational decision-making. (USAID 2016, p.3)

The decision tree (Figure I) presents different pathways to guide thinking about what an evaluation is intended to measure, quantify, or compare. This decision tree identifies types of evaluation (in red boxes) directly related to descriptions in Table I. The table describes types of program evaluation, with their respective methods of data collection and analysis, and relative costs.

# Decision Tree: Selecting the Right Kind of Evaluation for Your Question

Figure 1. Decision Tree for Program Evaluation Types



**Table I. Program Evaluation Types, Data Collection and Analysis Methods, and Costs**

Evaluation Type and Description	When to Carry Out	Types of Data and Information Produced	Data Collection Method	Illustrative Evaluation Questions for Nutrition Programs	Estimated Resource Requirements (time and \$-\$)\$) <sup>1</sup>
<b>Performance or outcome evaluation (non-experimental):<sup>2</sup></b> <ul style="list-style-type: none"> <li>Used to measure program or intervention effects in the target population by assessing progress in the outcomes or objectives that the program is to achieve</li> <li>Often incorporate before-after comparisons, but generally lack a counterfactual</li> </ul>	<ul style="list-style-type: none"> <li>During the design phase of a project, consider the performance evaluation(s) that will be undertaken. This is part of the preparation of a Project Monitoring, Evaluation, and Learning Plan.</li> <li>Carried out during and after program implementation or only after program implementation</li> </ul>	<ul style="list-style-type: none"> <li>Quantitative estimate of outcome changes in program area (can only assert contribution, not attribution, of changes to the program)</li> <li>Qualitative data on perceptions and behaviors</li> </ul>	<ul style="list-style-type: none"> <li>Beneficiary- or population-based quantitative surveys</li> <li>Qualitative interviews, focus groups</li> <li>Observation</li> <li>Participatory action research</li> <li>Monitoring and evaluation (M&amp;E) data review</li> <li>GIS/remote sensing analysis</li> <li>Process tracing<sup>3</sup></li> <li>Outcome mapping</li> <li>Outcome harvesting<sup>4</sup></li> </ul>	<p><b>General:</b> Did the program achieve its intended outcomes? Or, to what degree did the program achieve its intended outcomes?</p> <p><b>Nutrition:</b> What outcomes did participant households experience that could be due to receiving the activity's nutrition-sensitive agricultural training package? What challenges did households face in participating?</p>	<p>Moderate (\$\$) Time:</p> <ul style="list-style-type: none"> <li>~3–6 months for design and planning</li> <li>~6 months for data collection, analysis, writing (Note: this would repeat for each round of baseline/midline/endline data collection.)</li> </ul>

<sup>1</sup> A recent review of USAID program evaluations found that the cost of impact evaluations ranges from 1–32 percent of the project budget; for performance evaluations the range is 0.2–35 percent. The dollar symbols in this column indicate a relative (non-specific) cost range for each evaluation type.

<sup>2</sup> Performance indicators measure expected outputs and outcomes of strategies, projects, or activities based on the Mission's Results Framework or a project's or activity's logic model. In general, outputs are directly attributable to the program activities, while project outcomes represent results to which a given program contributes but for which it is not solely responsible. Performance evaluations may address descriptive, normative, and/or cause-and-effect questions: what a particular project or program has achieved (at any point during or after implementation); how it is being implemented; how it is perceived and valued; whether expected results are occurring; and other questions that are pertinent to design, management, and operational decision-making (USAID 2016).

<sup>3</sup> Process tracing: a case-based approach to causal inference which focuses on the use of clues within a case (causal-process observations) to adjudicate between alternative possible explanations, involving four types of causal tests (Davies 2016)

<sup>4</sup> Outcome harvesting: collects evidence of what has changed, and then, working backwards, determines whether and how an intervention has contributed to these changes (Wilson-Grau 2015)

Evaluation Type and Description	When to Carry Out	Types of Data and Information Produced	Data Collection Method	Illustrative Evaluation Questions for Nutrition Programs	Estimated Resource Requirements (time and \$-\$-\$) <sup>1</sup>
<b>Process evaluation (non-experimental):</b> Used to understand how program objectives or outcomes were achieved	Carried out during and /or after program implementation, or only after program implementation	<ul style="list-style-type: none"> <li>• Qualitative data on process</li> <li>• Quantitative outcome measurement (for reference)</li> </ul>	<ul style="list-style-type: none"> <li>• M&amp;E data review</li> <li>• Key informant interviews</li> <li>• Outcome harvesting</li> <li>• Process mapping</li> </ul>	<p><i>General:</i> How did the program achieve these outcomes? Or, specify certain outcomes (perhaps unexpected) to investigate.</p> <p><i>Nutrition:</i> How well did the program incorporate nutrition counseling into community health worker (CHW) services? How could that incorporation be improved? By how much did children's and women's dietary diversity improve among households receiving CHW-delivered nutrition counseling?</p>	Low (\$) Time: ~1 month to design and plan, 1 month to conduct, 1 month to analyze data and create report
<b>Impact evaluation (quasi-experimental):<sup>5</sup></b> <ul style="list-style-type: none"> <li>• Used to estimate program or intervention effectiveness in</li> </ul>	<ul style="list-style-type: none"> <li>• For each project, consideration will be given during the design phase to impact</li> </ul>	<ul style="list-style-type: none"> <li>• Quantitative estimates of the impact that can be attributed to the program</li> </ul>	<ul style="list-style-type: none"> <li>• Quantitative survey of treatment and comparison groups</li> </ul>	<p><i>General:</i> What outcomes can be attributed to the program or intervention?</p> <p><i>Nutrition:</i> What impact did the activity's nutrition-sensitive agricultural</p>	Moderate-high (\$-\$-\$-\$) Time: ~6–9 months for data collection, analysis, writing (Note: this would repeat for each)

<sup>5</sup> Impact evaluations measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change (USAID, 2016). Project design and the M&E plan supports the evaluation design. Part of the design and implementation process entails ensuring no one in the comparison group receives the intervention (or at least limiting the number of those who do).

<b>Evaluation Type and Description</b>	<b>When to Carry Out</b>	<b>Types of Data and Information Produced</b>	<b>Data Collection Method</b>	<b>Illustrative Evaluation Questions for Nutrition Programs</b>	<b>Estimated Resource Requirements (time and \$—\$\$\$)<sup>1</sup></b>
<ul style="list-style-type: none"> <li>achieving its ultimate goals</li> <li>Includes a comparison group</li> </ul>	<ul style="list-style-type: none"> <li>evaluation(s) that will be undertaken in some cases.</li> <li>During program implementation or after implementation (e.g., two years after).</li> </ul>	<ul style="list-style-type: none"> <li>Qualitative data on perceptions and behaviors</li> </ul>	<ul style="list-style-type: none"> <li>Qualitative interviews, focus groups</li> <li>Observation</li> <li>Participatory action research</li> <li>M&amp;E data review</li> <li>GIS/remote sensing analysis</li> </ul>	<p>training package have on children's and women's dietary diversity? Or, by how much did the intervention increase children's and women's diet diversity?</p> <p>How did participants increase consumption of nutrient-rich foods?</p>	round of baseline/midline/endline data collection.)
<b>Impact evaluation (experimental—Randomized Controlled Trial):<sup>6</sup></b> <ul style="list-style-type: none"> <li>Used to measure program or intervention effectiveness in achieving its ultimate goals</li> <li>Includes a rigorously defined counterfactual<sup>7</sup></li> </ul>	During program implementation (pre- and post-measurements)	<ul style="list-style-type: none"> <li>Quantitative impact estimates attributable to the program</li> <li>Qualitative data on perceptions and behaviors</li> </ul>	Quantitative survey of intervention and control groups <ul style="list-style-type: none"> <li>Qualitative interviews, focus groups</li> <li>Observation</li> <li>Participatory action research</li> <li>M&amp;E data review</li> <li>GIS/remote sensing analysis</li> </ul>	<p><b>General:</b> What outcomes can be attributed to the program or intervention?</p> <p><b>Nutrition:</b> What impact did the Activity's nutrition-sensitive agricultural training package have on children's and women's dietary diversity? Or, by how much did the intervention increase children's and women's diet diversity? How did participants increase consumption of nutrient-rich foods?</p>	High (\$\$\$) Time: ~6–12 months for data collection, analysis, writing (Note: repeat for each round of baseline/midline/endline data collection.)

<sup>6</sup> Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured (USAID 2016). Design at the beginning of the project and the M&E plan supports the evaluation design.

<sup>7</sup> Part of the design and implementation is to ensure no one in the counterfactual/control group receives the intervention.

<b>Evaluation Type and Description</b>	<b>When to Carry Out</b>	<b>Types of Data and Information Produced</b>	<b>Data Collection Method</b>	<b>Illustrative Evaluation Questions for Nutrition Programs</b>	<b>Estimated Resource Requirements (time and \$-\$-\$)</b> <sup>1</sup>
<p><b>Economic evaluation, cost analysis, cost-benefit analysis (non-experimental, could be a component of an evaluation with a quasi-experimental or experimental design)<sup>8</sup></b></p> <p>Used to identify, compare, quantify, and determine the value of the economic costs and benefits of a program or project<sup>9</sup></p>	Before, during, and/or after program implementation (see previous column)	<ul style="list-style-type: none"> <li>• Quantitative estimates of costs of program method options, compared to expected program outcomes</li> <li>• Quantitative and qualitative analysis of outcome changes in the program area given quantitative costs</li> </ul>	<ul style="list-style-type: none"> <li>• Cost-benefit</li> <li>• Cost effectiveness</li> <li>• Demand analysis</li> <li>• Alternative analysis</li> <li>• Risk and sensitivity analysis</li> </ul>	<p><b>General:</b> Which program approach is more cost-effective?</p> <p>What would be the cost of scaling up the program?<sup>10</sup></p> <p>What was the cost to produce an output (such as per person reached)?</p> <p><b>Nutrition:</b> What was the cost per nutrition outcome achieved (e.g., comparing costs for two interventions (social media messages vs. integrating nutrition counseling in well child visits) to their effectiveness in increasing the proportion of women consuming minimum dietary diversity?</p>	<p>Moderate (\$\$) as a standalone evaluation</p> <p>Low (\$) to add to another evaluation design</p> <p>Time: ~ 6–12 months for data collection, analysis, writing as a standalone evaluation; additional two months for data collection and analysis as part of a larger evaluation</p>
<p><b>Sustainability evaluation (non-experimental; could be a component of an evaluation with a</b></p>	At the end of a program or after a program has ended (e.g., two to five years later)	<ul style="list-style-type: none"> <li>• Using both quantitative and qualitative methods to examine program</li> </ul>	Quantitative and qualitative (may have a comparison group that did not receive the intervention)	<p><b>Intervention sustainability:</b></p> <ul style="list-style-type: none"> <li>• Will or has the intervention continue(d) after external funding has ended?</li> </ul>	<p>Moderate (\$\$) as a standalone evaluation;</p> <p>Low (\$) to add to another evaluation design</p>

<sup>8</sup> Refer to the [SEEMS-Nutrition common approach to measure the costs and benefits of multi-sectoral nutrition strategies](#). USAID Advancing Nutrition guidance on economic evaluation and costing tools forthcoming.

<sup>9</sup> Asian Development Bank 2017

<sup>10</sup> MEASURE Evaluation 2006

Evaluation Type and Description	When to Carry Out	Types of Data and Information Produced	Data Collection Method	Illustrative Evaluation Questions for Nutrition Programs	Estimated Resource Requirements (time and \$-\$)\$) <sup>1</sup>
<b>quasi-experimental or experimental design)</b> There are two types of sustainability evaluations: <ul style="list-style-type: none"> <li>sustainability of activities, which describes the likelihood of behavioral outcomes being maintained after the activity ends</li> <li>the level of sustained outcomes, which quantifies or otherwise describes the extent to which behavioral outcomes were maintained after an activity ended</li> </ul>		outcome data critically and in light of contextual factors <ul style="list-style-type: none"> <li>Using both quantitative and qualitative methods to re-examine program outcomes along with contextual factors</li> </ul>		<ul style="list-style-type: none"> <li>Was the impact of the program still evident after a given amount of time? (experimental)</li> </ul> <p>Sustained outcomes:</p> <ul style="list-style-type: none"> <li>Will the program outcomes be maintained without program inputs?</li> <li>Were the program outcomes maintained?</li> </ul> <p><i>Nutrition:</i> Will the observed increase in iron-folic acid (IFA) supplementation acceptance be maintained after external funding for distribution during antenatal care ceases?</p>	Time: ~6–12 months for data collection, analysis, writing as a standalone evaluation; additional three months for data collection and analysis as part of a larger evaluation

# Considerations for Conducting a Nutrition Program Evaluation

## I. Program Design and Evaluation: What Do You Want to Know? Who Can Find Out?

Several considerations guide decision making about evaluation and are relevant when designing programs. Consider how a program will be evaluated when the program is designed.

- When developing a Program Impact Pathway (PIP) and theoretical frameworks (e.g., Results Framework), also think about how progress and achievements will ultimately be judged. As the theoretical framework identifies what the program expects to accomplish, the evaluation determines if and to what extent those results were achieved (see Annex 2 for more on theoretical frameworks). Guiding questions at the formative stage may include—
  - Is it important to unequivocally attribute results to the program?
  - Or, is the main concern of the evaluation what was accomplished?

It may be enough to report accomplishments without 100 percent certainty that the program alone was responsible for them. Attribution is important when a new intervention or combination of interventions is tried or if there is an intention to scale up this program's interventions. In those cases, generating the best possible evidence about the program's contributions is warranted. However, if the program implements interventions with an established evidence base then it may be more important to document accomplishments and how the program was implemented. In that case, the evaluation is advised to put forward a plausibility argument<sup>11</sup> contending program attribution but will focus assessment on what and how results were achieved.

- Evaluators are responsible for managing and guaranteeing the quality assurance of the evaluation and its findings to ensure its utility for program implementers and stakeholders. According to the [USAID Evaluation Policy](#), evaluations completed by USAID and/or State grantees generally do not count toward policy compliance unless they are both a) specified in the grant agreement, with U.S. government funds allocated for that purpose and b) conducted by a party external to the grantee (USAID 2016). For USAID, these evaluations are identified in the registry as “internal evaluations.” Considerations of neutrality and partiality should be considered when selecting evaluators; critically, the evaluation team should be qualified and able to carry out a fair assessment of a project.

## 2. Identifying Evaluation Questions and Appropriate Methodologies

The evaluation questions guide the evaluation design. As the team considers the evaluation purpose, they will identify what they are most interested in knowing and learning about the program.

- The timing of evaluation planning influences the questions that can be asked. If the purpose is also to answer the question of attribution, then an evaluation must be planned alongside the program because it affects baseline data collection and implementation design. It is always best to design or at least identify some key components of an evaluation at the start of a program, but sometimes an evaluation is not or cannot be designed when a program is designed. In that case, the evaluation questions may be determined after program implementation. More flexible methodologies can examine what happened and how, but attribution for those observations cannot be made with certainty to the program or even plausibly argued in many cases.

<sup>11</sup> Plausibility argument: a set of propositions that makes the case for a certain conclusion

- Evaluations are carried out through a variety of methods, depending on the specific evaluation questions and the type of evaluation (Table 1). “Mixed methods” describes an approach incorporating both qualitative and quantitative assessments. This approach creates complementarity in the information uncovered. Using mixed methods provides the best information for program evaluations since the interest centers on discovering what happened and understanding why. In addition to matching the evaluation questions, the choice of methods will be affected by the time available for the evaluation and the availability of resources.
- Another consideration is the number of people needed to obtain the desired information. “Sampling” is how the number and type of people participating in the evaluation assessments are selected. The process differs by method and purpose, which is determined by the evaluation questions.
  - Quantitative methods: Sample size calculations are based on a parameter of interest (usually the estimated prevalence of an outcome) and the level of confidence you want in the result (i.e., how wide your margin of error can be). Since two groups are compared with an impact evaluation, it is useful to consult a statistician during the planning and analysis phases to ensure appropriate calculations and sample sizes.
  - Qualitative methods: Evaluators have more flexibility in choosing a purposive sample<sup>12</sup> of participants because these methods do not seek to produce statistical estimates. Evaluators seek “saturation,”—when the team has obtained all the information available about a particular topic from participants. This is evident when participants’ answers begin to repeat or corroborate others and may be obtained in as few as 6–12 interviews (using the same questions, with the same type of respondent) (Guest et al. 2006). Sampling and data collection methods should be made clear and available, allowing future researchers to implement them, and readers to determine the likelihood of bias (Goldberg et al. 2017). Box 2 calls out the role of Institutional Review Boards in program evaluation.

As the Center for Global Development has outlined, “quantitative evaluations make little sense without information about context, implementation, and the meaning people ascribe to programs. On the other hand, qualitative evaluations lack credibility without information (often quantitative) that corroborates or qualifies the information from interviews and focus group discussions”

(Goldberg et al. 2017).

## Box 2. Institutional Review Board (IRB) Considerations

Collecting information from people usually requires the approval of an Institutional Review Board (commonly called IRB) even if you are asking about seemingly neutral things such as their diet or where they buy food. It is appropriate to obtain IRB approval from the implementing organization and a scientific body in the country where the program was implemented. Appropriate measures should be taken to protect the privacy of participants and ensure the confidentiality of responses (Goldberg et al. 2017), including informed consent procedures; the IRB application will document those processes.

<sup>12</sup> Purposive sample: a type of non-probability sampling using at least one criterion to select the sample (Guest, Namey, and Mitchell 2013, p. 48)

### **3. Identifying Appropriate Indicators**

Indicators are measurements informed by quantitative data. For more resources related to standard nutrition indicators, see Annex 2.

- At the outcome level, there are standard references to inform program indicators, such as *Indicators for Assessing Infant and Young Child Feeding Practices: Definitions and Measurement Methods* from WHO and UNICEF (2021).
- Outputs and processes are measured more organically and aligned with the program design (e.g., the number of people trained or the number of coordination meetings held). Many processes are not measured with indicators; rather, evaluations describe them in terms of who participated, when something occurred, quality of products, trainings and services, etc.
- When an evaluation follows a program's theoretical framework, some indicators used by the evaluation team (mainly process and output, but sometimes also outcomes, such as behaviors) will match the indicators used to monitor those framework elements. Those monitoring data enable evaluators to conduct a trend analysis to examine changes over time.
- In cases where an evaluation team does not have a program theoretical framework to follow, they will create one post-hoc to guide the evaluation. Evaluators will ensure the indicators they select or develop answer the evaluation question(s) and that appropriate assessment methods are used to estimate those indicators. Evaluation reports should include data and methodological limitations so that readers can judge the validity of the reported findings.

### **4. Analyzing Data and Producing the Report**

In addition to quantitative and qualitative data analysis methods, evaluations may use specialized methods such as network analysis to learn about collaboration or cost effectiveness analysis. Regardless of the methods used, “triangulate” findings by using different methods and data sources to answer evaluation questions. This practice has the benefit of uncovering information that validates or contradicts other findings, with the latter spurring further investigation. It also enables evaluators to understand different aspects of a situation, process, output, or outcome.

- Development programs are by nature complex, making it unlikely that an evaluation using a single method or source of data can provide a full picture of what a program accomplished and what was learned. Complexity-aware monitoring and evaluation seeks to inform adaptive management approaches when cause and effect relationships are unclear or nonlinear, and rely on several data sources to ascertain what is happening. (See, for example, USAID 2018; MEASURE Evaluation 2017).
- Evaluation reports must follow the [U.S. Open Data Policy](#). Publishing data in non-proprietary, machine readable formats is helpful to evaluation teams and to peer reviewers and readers in considering the credibility of findings (Goldberg et al. 2017). In order to publish machine-readable datasets on Open Data, the participant informed consent form must indicate that aggregated data will be shared online for researchers to access. See examples of consent forms from published datasets (e.g., Feed the Future surveys).

### **5. Disseminating and Using Information**

A key component of planning program evaluations involves developing a dissemination plan and identifying audiences and goals for sharing and applying evaluation findings. The plan will inform the

design of the evaluation report and other products and ensure the evaluation furthers the program's goals.

- The evaluation team should plan ahead for the most appropriate and effective mechanism(s) for dissemination of the evaluation products, which may include: a live presentation or webinar for program implementers and stakeholders, posters for distribution to participating districts, or a published version of the report with an executive summary, or list of key takeaways. Program teams will want to consider how to package and explain the evaluation findings for different audiences.
- It is critical to engage program stakeholders throughout the evaluation process, from planning to dissemination, to ensure a successful evaluation, and dissemination and use of the findings. Stakeholders should also be at the forefront when planning the dissemination plan itself. Program evaluators should consider at which points stakeholders are consulted, including at the evaluation planning stage, within the evaluation itself, and to aid in dissemination of the findings. Stakeholders are then equipped to advise and implement related programs in the future. Evaluations are only valuable if seen as relevant and will be impactful in terms of informing policy or future programs if “owned” by prospective users (MEASURE Evaluation 2006).
- The following questions are useful for determining engagement of an appropriate set of stakeholders when proposing, designing, implementing, and reporting on evaluations (MEASURE Evaluation 2006):
  - Who needs to use the data? What questions are they seeking to answer?
  - Who has influence and resources that can aid this evaluation?
  - Who will be directly or indirectly affected by the outcome of this evaluation?
  - Who will support our plan? Who will oppose it? Why? How do we respond?
  - What roles can these people contribute to the process?
  - How can we best leverage their insights or assuage their objections?

# References

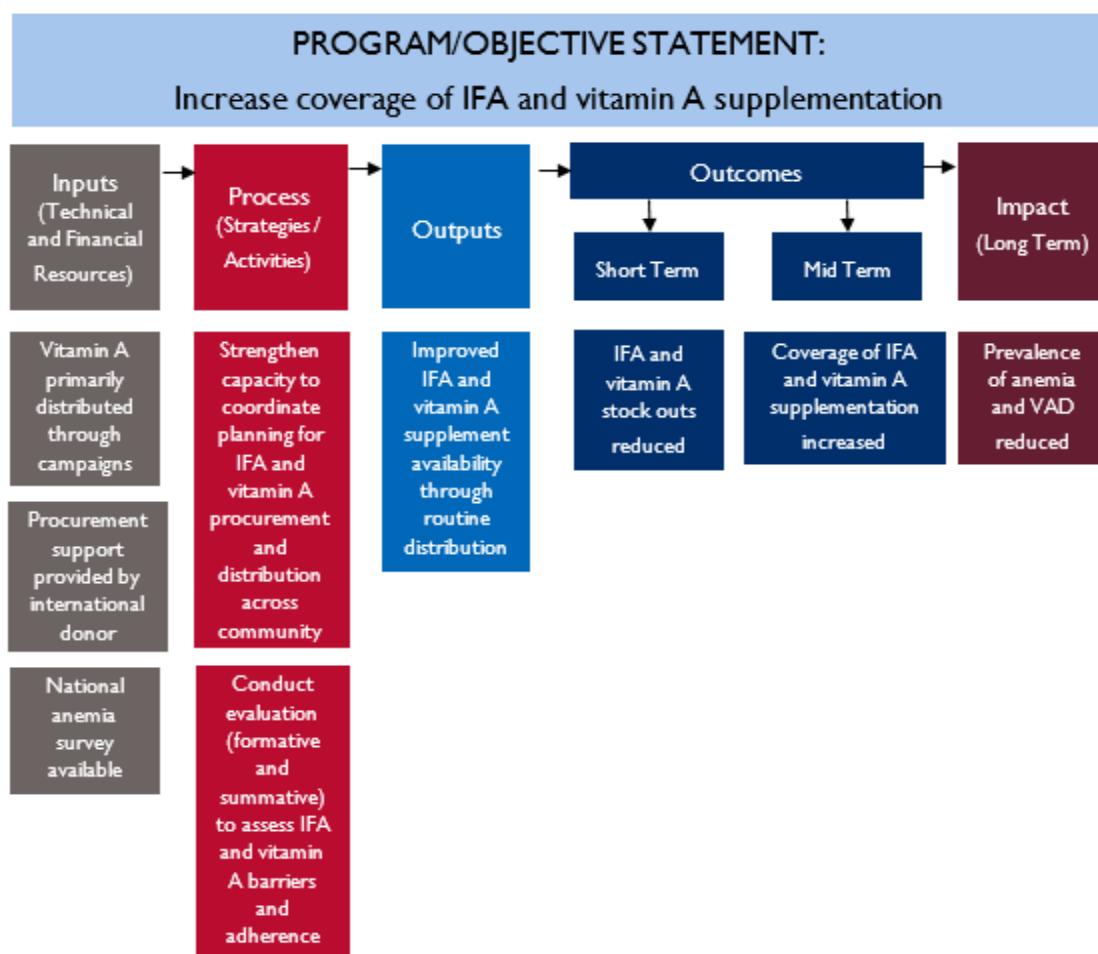
- Asian Development Bank. 2017, March. "Guidelines for the Economic Analysis of Projects." Accessed July 9, 2021. <https://www.adb.org/documents/guidelines-economic-analysis-projects>.
- BetterEvaluation. 2013, May. "Evaluation Costing." Accessed July 9, 2021. [https://www.betterevaluation.org/en/evaluation-options/calculate\\_evaluation\\_costs](https://www.betterevaluation.org/en/evaluation-options/calculate_evaluation_costs).
- Bill & Melinda Gates Foundation. 2011, January. "A Guide to Actionable Measurement." Accessed July 9, 2021. <https://intelligentmeasurement.net/2011/01/06/a-guide-to-actionable-measurement/>.
- Centers for Disease Control and Prevention. n.d. "Types of Evaluation: What Are the Most Common Types of Evaluation?" <https://www.cdc.gov/std/Program/pupestd/Types%20of%20Evaluation.pdf>.
- Davies, Rick. 2016, April. "Process Tracing," BetterEvaluation. Accessed April 22 2021. <https://www.betterevaluation.org/en/evaluation-options/processtracing>.
- Fehringer, J., E. Millar, and B. Wilkes. 2020. "Needs Assessment on Resources for Use of Routine and Other Secondary Data for Health Evaluation and Research." *Data for Impact Project*. Accessed April 10, 2021. <https://www.data4impactproject.org/publications/needs-assessment-on-resources-for-use-of-routine-and-other-secondary-data-for-health-evaluation-and-research/>.
- Goldberg Raifman, Julia, Felix Lam, Janeen Madan Keller, Alexander Radunsky and William Savedoff. 2017. "Evaluating Evaluations: Assessing the Quality of Aid Agency Evaluations in Global Health." Working Paper 461, Center for Global Development. Accessed April 10, 2021. <https://www.cgdev.org/publication/evaluating-evaluations-assessing-quality-aid-agency-evaluations-global-health>.
- Guest, G., A. Bunce, and L. Johnson. 2006. "How Many Interviews are Enough?: An Experiment with Data Saturation and Variability." *Field Methods*. 18(1):59–82. <https://doi.org/10.1177%2F1525822X05279903>
- Guest, Greg, Emily E. Namey, and Marilyn L. Mitchell. 2013. *Collecting Qualitative Data: A Field Manual for Applied Research*. Los Angeles: SAGE Publications, Inc.
- Hart L, B.S. Iskarpatyoti, M. Mandal, J.C. Thomas, 2016. "Evaluating Structural Interventions—Guidance for HIV Prevention Programs." *MEASURE Evaluation*. Accessed April 20 2021. <https://www.measureevaluation.org/resources/publications/tr-16-139>.
- International Initiative for Impact Evaluation. n.d., "Evidence Gap Maps." Accessed July 9, 2021. <https://www.3ieimpact.org/evidence-hub/evidence-gap-maps>
- Iskarpatyoti, Brittany, S. Beth Sutherland, Heidi W. Reynolds. 2017. "Getting to an Evaluation Plan: A Six-Step Process from Engagement to Evidence." *MEASURE Evaluation*. Accessed January 12, 2021. <https://www.measureevaluation.org/resources/publications/ms-17-124>.
- MEASURE Evaluation. 2017. "Complexity-Aware Methods." Accessed July 9, 2021. [https://www.measureevaluation.org/resources/publications/fs-17-217/at\\_download/document](https://www.measureevaluation.org/resources/publications/fs-17-217/at_download/document)
- MEASURE Evaluation. 2006. "Data Demand and Information Use in the Health Sector: A Conceptual Framework." Accessed July 9, 2021. <https://www.measureevaluation.org/resources/publications/ms-06-16a>.
- MEASURE Evaluation. 2019. "Data for Impact: What's the Cost of Evaluations and Other Surveys?" Accessed July 9, 2021. <https://www.measureevaluation.org/resources/publications/fs-19-361-d4i>.
- MEASURE Evaluation. 2016. "Types of Evaluations Conducted by MEASURE Evaluation." Accessed July 8, 2021. <https://www.measureevaluation.org/resources/publications/fs-15-158>.
- Silvestre, Eva. 2020. "A Practical Guide to Using Routine Data in Evaluation." Data for Impact Project. Accessed July 9, 2021, <https://www.data4impactproject.org/publications/a-practical-guide-to-using-routine-data-in-evaluation/>.
- Social Innovation Fund. n.d. "Evaluation Budgeting Quick Guide." Accessed July 9, 2021. <https://www.nationalservice.gov/sites/default/files/documents/SIF%20Evaluation%20Budgeting%20Quick%20Guide.pdf>.

- USAID Learning Lab. 2018, July. "Complexity-Aware Monitoring Discussion Note." Accessed July 9, 2021.  
<https://usaidlearninglab.org/library/complexity-aware-monitoring-discussion-note-brief>
- USAID Learning Lab. 2015, February. "Evaluation Toolkit." Accessed July 9, 2021.  
<https://usaidlearninglab.org/evaluation-toolkit>.
- USAID Learning Lab. 2019. "Project Evaluation Questions Template." Accessed July 9, 2021.  
<https://www.usaid.gov/project-starter/program-cycle/project-design/project-evaluation-overview/project-evaluation-questions-template>.
- USAID MOMENTUM. 2020, December. "A Guide to Complexity-Aware Monitoring Approaches for MOMENTUM Projects." Accessed July 9, 2021. <https://usaidmomentum.org/resource/a-guide-to-complexity-aware-monitoring-approaches-for-momentum-projects/>.
- UNICEF. 2018, April. "Revised Evaluation Policy of UNICEF 2018." Accessed July 9, 2021.  
<https://www.unicef.org/evaluation/documents/revised-evaluation-policy-unicef-2018>.
- UNICEF. 2017, June. "UNICEF-Adapted UNEG Evaluation Reports Standards." Accessed July 9, 2021.,  
<https://www.unicef.org/evaluation/documents/unicef-adapted-uneq-evaluation-reports-standards>.
- USAID. 2013, September. "Technical Note: Impact Evaluations." Accessed July 8, 2021.  
[https://www.usaid.gov/sites/default/files/documents/1870/IE\\_Technical\\_Note\\_2013\\_0903\\_Final.pdf](https://www.usaid.gov/sites/default/files/documents/1870/IE_Technical_Note_2013_0903_Final.pdf).
- USAID. 2016, January. "USAID Evaluation Policy." Accessed July 9, 2021. <https://www.usaid.gov/evaluation/policy>.
- WHO and UNICEF. 2021. "Indicators for Assessing Infant and Young Child Feeding Practices: Definitions and Measurement Methods." Accessed July 13, 2021. <https://www.who.int/publications/item/9789240018389>
- Wilson-Grau, R., 2015. "Outcome Harvesting." *BetterEvaluation*, Accessed July 9, 2021.  
[http://betterevaluation.org/en/plan/approach/outcome\\_harvesting](http://betterevaluation.org/en/plan/approach/outcome_harvesting)
- World Bank Group. 2019. "World Bank Group Evaluation Principles." Accessed July 9, 2021.  
<https://ieg.worldbankgroup.org/sites/default/files/Data/reports/WorldBankEvaluationPrinciples.pdf>.
- World Bank Group. 2020, November. "Guidance Notes on Evaluation Approaches and Methods in Development." Accessed July 9, 2021. <http://ieg.worldbankgroup.org/evaluation-international-development/guidance-notes-evaluation-approaches-and-methods-development>.
- World Health Organization, UNICEF, USAID, AED, UC, Davis, IFPRI. 2010. "Indicators for Assessing Infant and Young Child Feeding Practices." Accessed July 9, 2021.  
<https://www.who.int/nutrition/publications/infantfeeding/9789241599290/en/>

# Annex I. Additional Resources

Theoretical frameworks, or theories of change, explain program logic by illustrating the progression of effort and expected consequences through inputs, processes, outputs, outcomes, and impact. Different types of theoretical frameworks include conceptual models on one end of the spectrum (guiding planning but not typically measured) to results frameworks on the other (guides measurement), with PIPs, logic models, among others in between. In Figure 2 below, we use an example of a theoretical framework to explain the evaluation considerations for a program to increase IFA and vitamin A distribution. Evaluations are easier to assess if evaluators are explicit about the theoretical framework that underpins their analysis (Golberg et al. 2017), and ideally, the evaluator using the same framework that was used to design the program and not design one just for the evaluation.

Figure 2. Example of Theory of Change for a Nutrition Program<sup>13</sup>



For a performance evaluation of the program illustrated in Figure 2, the evaluation team would look at the anticipated outcomes and create an assessment plan appropriate to measure each one. Often, a population-based survey is used to assess outcomes, because it provides information about indicators of interest (such as coverage of vitamin A distribution) in a population. A performance evaluation may also

<sup>13</sup> IFA = iron-folic acid, VAD = vitamin A deficiency

be used to assess outputs. The evaluation team would want to know how many routine distribution events occurred and how many people were served.

Performance and other types of evaluations (described in Table 1) are also focused on processes—documenting how outputs and outcomes were achieved. Using the example in Figure 2, the evaluation team would want to review capacity strengthening assessments and reports from the barrier analysis. Then, interview capacity strengthening activity participants and program staff to understand how the capacity strengthening work and barrier analysis, respectively, affected program operations and service provision to produce the observed outputs and outcomes. Inputs are not typically assessed, but information about inputs is collected and presented as part of contextual information about the project. Impacts are generally not assessed. They represent overall goals but are usually not measured to assess performance, because they are long-term and high-level results, such as anemia prevalence or stunting, to which many things beyond the project contribute. The impacts in a theoretical framework are not to be confused with impact evaluation, which rigorously assesses program performance, “by estimating what would have happened in the absence of the program or aspect of the program” (GAO 2021) and can attribute findings to the program. (see Table 1).

In this example, quantitative methods would provide an estimate of coverage of vitamin A distribution in the form of a percentage of respondents who report receiving vitamin A, and qualitative methods would provide information about how vitamin A was distributed and the challenges faced by providers and recipients. This information informs a program evaluation and provides learning for future programs.

## Theoretical Frameworks

USAID Advancing Nutrition. 2018, September. “Theory of Change Training Curriculum.” Accessed July 9, 2021. <https://www.advancingnutrition.org/resources/theory-change-training-curriculum>

USAID Learning Lab. 2021. “Impact Evaluation Decision.” Accessed July 9, 2021. <https://www.usaid.gov/project-starter/program-cycle/project-design/project-evaluation-overview/impact-evaluation-decision>.

USAID Learning Lab. 2021. “Impact Evaluation Decision Tree.” Accessed July 9, 2021. <https://www.usaid.gov/project-starter/program-cycle/project-design/project-evaluation-overview/impact-evaluation-decision-tree>.

USAID Learning Lab. 2021. “Impact Evaluation Designs.” Accessed July 9, 2021. <https://www.usaid.gov/project-starter/program-cycle/project-monitor-evaluation-plan/monitor-evaluation-plan-evaluation-component/impact-evaluation-designs>.

USAID Learning Lab. 2016, March. “What is This Thing Called ‘Theory of Change?’” Accessed July 9, 2021. <https://usaidlearninglab.org/lab-notes/what-thing-called-theory-change>

USAID Learning Lab. 2017, September. “How-To Note: Developing a Project Logic Model,” Accessed July 9, 2021. <https://usaidlearninglab.org/library/how-note-developing-project-logic-model-and-its-associated-theory-change>

USAID. 2019, April. “Results Framework.” Accessed July 9, 2021. <https://www.usaid.gov/project-starter/program-cycle/cdcs/strategy-and-results-framework/results-framework>

USAID. 2013, July. “Technical Note: Developing Results Frameworks.” Accessed July 9, 2021. [https://www.usaid.gov/sites/default/files/documents/1865\\_508\\_RF\\_Technical\\_Note\\_Final\\_2013\\_0722.pdf](https://www.usaid.gov/sites/default/files/documents/1865_508_RF_Technical_Note_Final_2013_0722.pdf)

## Standard Nutrition Indicators

Centers for Disease Control and Prevention, 2016, March. “Nutritional Status Indicators.” Accessed July 9, 2021. [https://www.cdc.gov/nccdphp/dnpao/growthcharts/training/overview/page5\\_1.html](https://www.cdc.gov/nccdphp/dnpao/growthcharts/training/overview/page5_1.html)

Global Nutrition Report. 2020, May, “Appendix I: Nutrition Indicators.” Accessed July 9, 2021. <https://globalnutritionreport.org/reports/2020-global-nutrition-report/appendix-i-nutrition-indicators/>

United Nations System Standing Committee on Nutrition. 2015, February. "Nutrition Targets and Indicators for the Post-2015 Sustainable Development Goals." Accessed July 9, 2021.  
[https://www.unscn.org/files/Publications/Post\\_2015\\_Nutrition\\_Targets\\_and\\_Indicators\\_final\\_March\\_2015\\_website\\_.pdf](https://www.unscn.org/files/Publications/Post_2015_Nutrition_Targets_and_Indicators_final_March_2015_website_.pdf)

USAID Learning Lab. n.d., "Nutrition Indicator Reference Sheets." Accessed July 9, 2021.  
<https://usaidlearninglab.org/sites/default/files/resource/files/Nutrition%20Indicator%20Reference%20Sheets.s2.pdf>

## Key Terms

The General Accounting Office (GAO) released key terms and concepts for program evaluation (2021), including creating conditions for quality evaluations (such as conducting data reviews), establishing a theory of program change, engaging stakeholders, creating an evaluation plan, and conducting a capacity assessment, which we discuss beginning with theoretical frameworks.

Government Accountability Office. 2021, March. "Program Evaluation: Key Concepts and Terms." Accessed July 9, 2021. <https://www.gao.gov/products/gao-21-404sp>



### USAID ADVANCING NUTRITION

Implemented by:  
JSI Research & Training Institute, Inc.  
2733 Crystal Drive  
4<sup>th</sup> Floor  
Arlington, VA 22202

Phone: 703-528-7474  
Email: [info@advancingnutrition.org](mailto:info@advancingnutrition.org)  
Web: [advancingnutrition.org](http://advancingnutrition.org)

USAID Advancing Nutrition is the Agency's flagship multi-sectoral nutrition project, addressing the root causes of malnutrition to save lives and enhance long-term health and development.

---

This document was produced for the U. S. Agency for International Development. It was prepared under the terms of contract 7200AA18C00070 awarded to JSI Research & Training Institute, Inc. The contents are the responsibility of JSI and do not necessarily reflect the views of USAID or the U.S. Government.