



Technical Note

Evaluation Approaches, Methods and Data Collection Tools for Decentralized Evaluations

1	2	3	4
Introduction	Definitions	Key Evaluation Approaches	Main examples of Evaluation Methods
5	6	7	8
Main examples of data collection tools	WFP DEQAS requirements in relation to evaluation methods and data collection and analysis methods/ tools	Responsibilities	Further reading

1. Introduction

1. The purpose of this Technical Note is to provide Evaluation Managers of Decentralized Evaluations with a short introduction on evaluation approaches, methods and data collection tools to help them define the overall methodological approach in the evaluation ToRs as well as review critically the methods elaborated by the evaluation team at inception phase. This Technical Note does not cover centralized evaluations for which specific requirements are defined in their respective CEQAS. Similarly, it does not cover impact evaluations, for which the evaluation questions require a rigorous measurement of impact and for which a particular type of evaluation approach, experimental or quasi-experimental, is needed and requires specialized expertise. COs interested in conducting an impact evaluation should consult the [Impact Evaluation Decision Guide](#) and reach out to OEV Impact Evaluation Team for further support.
2. This Technical Note serves as an entry point to the vast range of literature and guidance on evaluation approaches, methods and tools available, and includes references to selected external resources for those who want to learn more.

2. Definitions

3. The main terms used in this Technical Note are defined as follows:
 - An **evaluation approach** (also referred as evaluation design) is the overall framework for the evaluation and sets out what questions are selected and how the evaluation will be conducted. Examples of evaluation approaches include experimental, quasi-experimental and non-experimental evaluations; they can also be categorized as formative, summative or developmental evaluations. An evaluation approach usually combines more than one method to answer the evaluation questions.
 - **Evaluation methods** “provide what information should be collected, from which source(s) it should be collected, for what purpose it should be collected and how the collected data will be analyzed in order to answer the evaluation questions. [...] The methodology must also indicate, in analyzing data, what benchmarks will be used in making the assessment for each evaluation criteria or question.”¹
 - **Data collection tools** are ways or channels for collecting the data required to answer the evaluation questions; they can include qualitative or quantitative techniques, such as key informant interview, document review or surveys.

3. Key Evaluation Approaches

4. Table 1 outlines a range of evaluation approaches and summarizes their differences based on their expected use and the specific questions they aim to address. Although other approaches exist, this Technical Note does not provide detailed guidance on quasi-experimental and experimental approaches as they are primarily used in the context of impact evaluations. For more information on these, reach out to OEV Impact Evaluation Team.

Table 1: Broad categories of evaluation approaches

Approach	Uses and types of questions	Description
Experimental versus non-experimental and quasi-experimental		
Non-experimental	<p>Used in all evaluation types that are not impact evaluations.</p> <p>Primarily used to answer descriptive/normative questions, such as ‘Did the intervention achieve its intended results?’</p>	<p>Apply to descriptive/normative questions and suggest a link through contribution, by building an argument by plausible association between the intervention and observed changes (for example: following five years of project implementation, what changes are observed that could be connected to WFP’s engagement?). This approach heavily relies on documentation that shows the logic of the design, the theory of change or change pathways, then implementation processes and recorded results. If such documentation is weak evaluators have to find ways of reconstructing them, for example through key stakeholder interviews. Control or comparison groups are not required.</p> <p>Can be implemented without baselines or monitoring data (for example through a case study design), although strong and regular monitoring data help to buttress progression toward change over time. A clear evaluation approach and matrix are required that set out how each question is answered (with what methods and data) so that a ‘line of sight’ can be shown from the questions to the end results.</p>

¹ UNEG norms and standards for Evaluation (2016).

Approach	Uses and types of questions	Description
		Although non-experimental evaluation methods can seem intuitively the easiest group of approaches, they require significant analytical and qualitative methods expertise to implement well.
Quasi-experimental (e.g. pretest and posttest for a treated and comparison group)	Used in impact evaluations to answer cause and effect questions (For example, 'What is the impact of WFP's nutrition activity on nutritional indicators in the target population?')	<p>This approach requires the same level of technical skills as experimental approaches in terms of identification and analysis, however, they are significantly more demanding in terms of data requirements. It addresses cause and effect questions and is used in impact evaluation. Require the identification of a <i>counterfactual</i> and the identification of a 'comparison' group to measure this. The counterfactual is what would have happened in the absence of the intervention. The comparison group of individuals will be identified through statistical means to ensure it resembles as closely as possible the group receiving the intervention in order for comparisons to be made.</p> <p>It can be applied once programme implementation has begun, although it requires existing data preceding the intervention to demonstrate parallel trends between intervention and comparison groups and establish a credible counterfactual ruling out alternative explanations.</p>
Experimental (e.g. Randomized Control Trial (RCT))	Used in impact evaluations to answer cause and effect questions (see example above).	<p>Address cause and effect questions and require the identification of a counterfactual through randomly assigning potential beneficiary population to a control group or to a treatment group at the time of the intervention design. The control group does not receive or receive a different set of activities with different timelines, compared to the intervention group. The intervention and control group will otherwise be similar and have the same attributes. The random assignment allows the assessment of the impact of the intervention with high levels of validity and confidence.</p> <p>Requires high level of resources (financial and human), stability of implementation and intervention, and specialized skills. The evaluation must be designed alongside the intervention. Nobody should be denied access to the intervention or treatment purely for the purposes of evaluation, which means that this design is only suitable when this condition can be met (for example, there is a funding or timing constraint which means that the full eligible population cannot be reached at once).</p>
Formative versus summative and developmental		
Summative evaluation	Used mainly for accountability but does not preclude from having also a learning objective, after the program's completion / at the end of a programme cycle,	Usually outcome-focused and ex-post, measures outcomes against pre-determined goals and frameworks at the end of an intervention, with the view to both meet accountability requirements and inform decision making. It looks at how effectively the program made the desired change happen and how the program changed the lives of program participants and whether these changes are sustainable, therefore judging the merit or worth of the intervention at its conclusion.

Approach	Uses and types of questions	Description
	helps answering questions such as: "Was the programme relevant to the needs of the targeted populations? What are its outcomes?"	
Formative evaluation	Used to enhance learning on ongoing programs, helps answering questions such as: "What does and does not work? What are current strengths and weaknesses?"	Usually takes place before or during a programme's implementation to inform/ improve the design and performance. The results provide an early assessment of whether desired outcomes are likely to be achieved. They assist identifying and correcting implementation problems for the continuous improvement of the programme being evaluated. A formative evaluation assumes there is a pre-existing model and seeks to identify tweaks or improvements to this model. It looks at if a programme works and what factors (internal and external; enabling or hindering) come into play. Requires more qualitative methods of inquiry, with open questions.
Developmental evaluation	Used to assess developing or emerging concepts, ideas and initiatives, helps answering questions such as: "Does the programme consider external events and limitations into account? What are general patterns across programmes?"	Focuses on adaptive capacities in complex dynamic systems, in contexts that are rapidly changing, uncertain and turbulent e.g. innovations, crises. This allows the evaluation to maintain a broader perspective, contributing to a response/ intervention that is being developed while questioning the original assumptions behind the intervention with a high degree of flexibility, openness and receptiveness to adapt the evaluation process where needed. It is based on systems thinking and allows to focus on emergence, similar to the design of innovation processes. Developmental evaluations are situated within the wider context of utilization-focused evaluation, providing regular evaluative inputs to promote ongoing learning.
Other general / broad approaches		
Utilization-focused evaluation	Used to enhance the utility of the findings for its primary intended users, asking: "Who is going to use the evaluation? What needs to be done to make it as useful for them as possible?"	Based on the principle that an evaluation should be judged on its usefulness to its intended users. It is planned and implemented in a way that increases the likelihood of the findings being used, identifying the primary users and uses and ensuring they are proactively engaged throughout the evaluation. Primary users, identified at the start, can explain how they intend to use the evaluation findings before decisions are made on the evaluation questions or methodologies. It requires more flexibility and time to fully engage with different users at different stages. It is compatible with other approaches.

Approach	Uses and types of questions	Description
Participatory evaluation	Used to empower beneficiaries/participants to better analyze their situation and produce more reliable findings	Actively involves the stakeholders of a programme or policy in the design and implementation of the evaluation at any stage of the process. It identifies early on which stakeholders should be involved and why, and what kind of participation is feasible. It allows participants to identify their own objectives and/or indicators of change. Several different viewpoints can be included in the findings as opposed to an overall consensus and recommendations are actively disseminated with beneficiaries. It can be used alone or in combination with other approaches.
Theory-based evaluation	Used when there is some predicted change to assess, helps answering questions on what worked, why and how	Based on an explicit theory of change or logic model of the programme and attempts to assess change at each stage of the theory to test the linkages (assumptions) between different levels of change. The theory of change developed is used to guide the evaluation and collect evidence to establish whether and/or how the intervention produced the desired changes
Realist evaluation	Used to improve understanding about how development interventions work in different contexts, helps answering questions on what works, for whom, in what respects, to what extent, in what contexts, and how?	Based on the assumption that programmes work under certain conditions and are influenced by the way that different stakeholders respond to them. It is focused on causation, assessing which initiatives contribute to different results and how. It is used to test and refine the programme theory and to determine whether and how the programme worked in a particular setting and its outcomes. It requires an in-depth understanding of how interventions work for different groups.

- The evaluation approach or design should bring together in an integrated and coherent manner a diverse set of evaluation methods to answer the evaluation questions selected, the scope and objectives, considering important elements such as the availability of a theory of change as well as the overall evaluation context (time, data and budget constraints). Simple evaluation questions can have relatively simple designs, with a limited range of methods used in a straightforward fashion. More complex questions demand more elaborate evaluation methods for collecting and analyzing the evaluation data. Either way, the driver behind the decisions on design and methods is the set of evaluation questions to be answered. The [Technical Note on Evaluation Criteria and Questions](#) provides more information on this.
- The main analytical framework for an evaluation, the **evaluation matrix**, maps the evaluation questions against the evaluation methods, indicators or/and lines of inquiry, data collection tools and sources of information. To develop the matrix, the evaluation team should list the evaluation questions, break them down into sub-questions and for each one, identify what data will be collected to answer the questions, which data collection methods will be used, from which sources, how the data will be analyzed and assess the strength of the evidence. The evaluation team will then use it to guide the analysis, ensure that all data collected is analyzed and triangulated and identify any evidence gap. For more information, see the [Technical Note on Evaluation Matrix](#).
- Good evaluation questions are ideally drawn from the intervention [Theory of Change \(ToC\)](#) to ensure specificity to the context and the intervention. If no ToC exists, or was superseded by

events, the evaluation team, with the support of programme colleagues and other stakeholders may reconstruct retrospectively a ToC to describe more in detail how the intervention is understood to produce a series of results. The [Technical Note on Logical Models / Theory of Change](#) provides more information on this.

4. Main examples of Evaluation Methods

8. Methods in evaluation are applied social science research methods; they relate to the process of how data will be gathered overall, and how it will be analyzed in order to answer evaluation questions. Table 2 summarizes some of the most common non-experimental evaluation methods, but there are many others.

Table 2: Main examples of evaluation methods

Evaluation method	Description
<u>Contribution analysis</u>	<ul style="list-style-type: none"> • It is a methodology used to identify the contribution a development intervention has made to a change or set of changes. The aim is to produce a credible, evidence-based narrative of contribution, rather than to produce conclusive proof. Contribution analysis is based on a recognition that it is difficult to prove attribution for many development interventions, so it assesses the causal connections, reducing uncertainty about the contribution the intervention is making to the observed results. Contribution analysis is designed to be used alongside theories of change that explicitly set out how change is, or was, supposed to happen at different levels; and compares reality with the theory. It is based on a clear theory of change and should not be used to develop alternative theories of change. It does not need baselines and is useful when looking at the replication or expansion of a programme.
<u>Qualitative comparative analysis</u>	<ul style="list-style-type: none"> • It is a methodology that enables the analysis of multiple cases in complex situations and can help explain why change happens in some cases but not others. QCA uses both quantitative and qualitative analysis to systematically and transparently generate findings across multiple case studies, comparing them. It requires in-depth knowledge of cases (often part of qualitative analysis) but is also capable of generating findings that can be generalized across wider populations (quantitative analysis). It can either require the collection of new data or build upon data that has been collected previously.
<u>Process tracing</u> <u>Most Significant Change</u>	<ul style="list-style-type: none"> • It is a qualitative analysis methodology used to see if results are consistent with the program theory and if alternative explanations can be ruled out. It is aimed to establish whether, and how, a potential cause or causes influenced a specified change or set of changes. It focuses on the use of clues / a set of formal tests within a case to examine the strength of evidence and adjudicate between alternative possible explanations. It assesses causation by testing alternative ideas about how change might have come about in order to confirm some hypothesis and/or eliminate others. • It is a qualitative analysis methodology used to systematically gather information from participants on the changes experienced as a result of the intervention. These stories of change are prioritized and assessed with participation from the same stakeholder groups, who will decide which they consider to be the most significant stories of all and check their accuracy. The method is particularly helpful for examining how and why change happens, what factors support change and which ones obstruct it, and which contexts best support change. It can be useful in surfacing (unexpected) changes which weren't necessarily considered when developing the results framework / TOC,

Evaluation method	Description
	and when there are no pre-set defined indicators. It helps to understand changes through the eyes of different stakeholders.
<u>Outcome harvesting/ evidencing</u>	<ul style="list-style-type: none"> This methodology has action learning at its core. It collects evidence of what has changed (“outcomes”). Then, working backwards, it determines whether and how an intervention has contributed to these changes. This method involves the key stakeholders as “change markers”. It is used in complex situations when relations of cause and effect are not fully understood and when the focus is mostly on outcomes rather than activities or output. It is not designed to assess whether or not activities were carried out according to plan, but rather in contexts where plans need to be constantly modified over time. There is a risk of bias towards outcomes that are easy to identify, and away from those that are more difficult to measure.
<u>Social network analysis</u>	<ul style="list-style-type: none"> This methodology is designed to help map and analyze social networks, to identify and analyze the relationships within and between different actors, which can be either individuals, groups or organizations. The results are presented with complex diagrams and maps of network. It is useful for assessing policy influencing and mobilization, where the work is carried out through partnerships and coalitions; and to map and analyze knowledge networks and communities of practice.
<u>Social mapping</u>	<ul style="list-style-type: none"> Used to present information on community layout, infrastructure, demography, ethnic or language groups, health patterns, wealth and other community issues/resources, climate/natural resource patterns, seasonal calendars, etc. It is made by the participants, not the evaluators, and it is not drawn to scale. It shows what participants believe to be relevant and important for them, reflecting their perceptions of their reality.

5. Main examples of data collection tools

9. WFP decentralized evaluations use both primary and secondary data to answer evaluation questions, but usually tend to mostly rely on secondary data. Primary data are those that the evaluation itself collects, while secondary data come from other sources – WFP country offices, other UN Agencies, non-governmental organizations or governments. Both primary and secondary data can be either qualitative or quantitative. A given tool might collect both qualitative and quantitative data. For example, through an interview, a key informant can be asked for his opinion or experience of something, putting a numerical rating on it (e.g. “On a scale of one to five, where one is never and five is every day, how often in the last month did you feel hunger that you could not satisfy?”). In that way, this question asking about something qualitative – his experience of hunger – becomes quantitative, with a number on a scale. A quantitative survey might also include some open-ended qualitative questions.
10. Data collection for WFP decentralized evaluations typically is done through individual interviews, group interviews, observation/image recording and mini surveys. In some cases, they involve more comprehensive surveys. When access restrictions cannot be overcome, possibilities also exist to collect data remotely.
11. The balance between types of data will differ according to the nature of the evaluation questions to be answered. Table 3 shows some of the most common data collection tools.

Table 3: Most common data collection tools

Tool	Description	Quantitative	Qualitative
<u>Surveys</u>	<ul style="list-style-type: none"> • Structured, with a medium/high level volume of quantitative and qualitative information • Usually closed-ended questions with a pre-defined answer, like: Did you have breakfast this morning? Yes/No/Don't know. Open-ended questions can sometimes be used • Surveys can record the status of something, like: Are there cafeteria facilities? Complete/Partial/None/Don't know • When needed and feasible, these could be undertaken through SMS/text messages on mobile phones or interactive voice response (IVR), among others 	✓	
<u>Key informant interviews</u>	<ul style="list-style-type: none"> • Qualitative interviews with people who know details about the activity, like WFP, government, UN agencies, donors, community leaders. They usually include semi-structured open-ended questions. When needed and feasible, these can also be conducted through live calls, using Computer-Assisted Telephone Interviewing (CATI) 		✓
<u>Focus group discussions</u>	<ul style="list-style-type: none"> • Group interviews with structure and space for open answers • Gathers opinions and views, local perspectives and experiences (qualitative data) • Homogenous groups of 8-12 people (men and women in different groups, for example) are recommended 		✓
<u>Observation</u>	<ul style="list-style-type: none"> • A coded checklist to record observable events or behavior like under-scooping rations, or distribution preferences (quantitative) • An open-ended recording / visual intake of an experience, such as participants' use of vouchers, or food distribution (qualitative) • When needed and feasible, these can also be conducted through remote sensing/ satellite imagery or geospatial technology. 	✓	✓
<u>Self-reported checklists / Report Cards</u>	<ul style="list-style-type: none"> • Respondent perspectives: time spent working or cooking, taking produce to market, or how they used their vouchers • Can collect quantitative information (e.g. frequency, satisfaction) or qualitative information (descriptions of the experience) • Usually relies on informants being literate 	✓	✓
<u>Case studies/ Stories</u>	<ul style="list-style-type: none"> • Collect narratives from individuals such as recipients of assistance about their experiences, usually in great depth, and examining the story from a range of different angles (qualitative data) 		✓
<u>Diaries, journals</u>	<ul style="list-style-type: none"> • Gathers in-depth qualitative information about life events, over a long period of time • Can be applied to monitor such interventions like nutrition support, to examine eating patterns by family member 		✓

Tool	Description	Quantitative	Qualitative
<u>Review of secondary data</u>	<ul style="list-style-type: none"> Quantitative and/or qualitative data that has already been collected (and possibly analyzed) by WFP or other agencies. E.g. anthropometric data in national household panel surveys; retail price data from WFP Market Assessments. 	✓	✓
<u>Desk review</u>	<ul style="list-style-type: none"> Documentation analysis and literature scanning of internal and external documents to gather background information and determine the implementation of a programme 		✓

6. WFP DEQAS requirements in relation to evaluation methods and data collection and analysis methods/ tools

13. As per UNEG standards 4.5 “Evaluation methodologies must be sufficiently rigorous such that the evaluation responds to the scope and objectives, is designed to answer evaluation questions and leads to a complete, fair and unbiased assessment.” There is no “right” method to answer any evaluation question. Evaluation Managers should work with the evaluation team to find a strong evaluation approach that can give WFP answers to the evaluation questions they have. The choice of evaluation methods and data collection tools, goes hand in hand with the evaluation questions.
14. A high-quality evaluation approach ensures rigor in the evaluation process and produces reliable data and findings. It contributes to the credibility and completeness of the evaluation. A good test for the reliability of the data collection methods and data is to ask ourselves a key question: if another evaluation team followed this evaluation approach, would they get the same results and draw the same conclusions?
15. **Mixed methods and triangulation:** A question is likely to have different answers depending on whom you ask. For this reason, **WFP recommends using multiple and mixed methods** in its evaluations. Incorporating multiple and mixed methods into an evaluation results in a stronger, more complete evaluation than relying on only one method. A mixed method design systematically integrates two or more methods, usually drawing on both quantitative and qualitative data.
16. For example, interview responses from programme staff can be cross-checked with focus group data from targeted communities. These opinions and responses often differ because the people answering the questions have different perspectives. Another source, such as observations made during site visits or documentary data, may help to resolve any differences. This is called **triangulation** and gives stronger evidence on which to base evaluation conclusions. Where qualitative and quantitative data are also triangulated for a given question, you can generally have more confidence in the findings and the evaluation will have more credibility. The goal of triangulating with various data sources is to counteract the interests or biases found in any one data source. For example, farmers who discuss their contribution to a school meals programme may not recognize nutritional gaps in their produce, but the cooks in the school, or the school canteen records about what was provided and to how many children may reveal the need for more inputs. Later, when the evaluation team analyzes the data and writes the report, such comparative information should inform evaluation findings.

17. In addition to simply comparing data on the same questions, triangulation can also involve weighting the different evidence. The evaluation team will need to rate each source in a systematic way, and some sources would therefore have more weight in analysis. Such weighting should be transparent, and the evaluation team should discuss how they would do so in the inception report.
18. Triangulation is typically carried out by the evaluation team at the reporting stage once the data collection has been completed. See Table 4 for different strategies for incorporating mixed methods into an evaluation.

Table 4: Strategies for using Multiple or Mixed Methods in Evaluation

Strategies	Examples
Different methods are used to answer different evaluation questions within one evaluation.	One evaluation question is answered through a survey (<i>do students report better attendance when school meals are provided?</i>), while another is answered through stakeholder interviews (<i>How can the programme be made sustainable?</i>).
Different methods are used to triangulate answers to the same question.	Stakeholder surveys might be used with beneficiaries to find out whether humanitarian assistance reached them. This is combined with records from GPS mapping of delivery vehicles and WFP monitoring records, and interviews with project staff.
Different methods are used to answer different <i>parts</i> of the same question.	A quantitative survey is used to identify if an intervention improved the well-being of intended recipients, while qualitative data is used to assess the <i>quality</i> of the improvement.
Different methods are used at different times in the evaluation process.	Qualitative data from stakeholder interviews at preparation stage is used to identify the most useful evaluation questions, which are then explored further with a targeted survey at inception stage and then focus groups, observations and case studies at data collection stage to add more depth to the findings. Data could be gathered in a sequential manner, using one type of data to inform the collection of other type of data through an integrated evaluation approach. This allows to combine different methods during the evaluation process to provide more insightful understandings, instead of gathering all the data at the same time independently and then combining them only at the end for interpretation with limited time for further enriching and triangulating the findings.

19. **Gender and equity considerations:** To ensure the evaluation approach is inclusive of diverse groups, the EM and the evaluation team should think critically about who might not get a chance to share their perspective. It is the responsibility of both the EM and the evaluation team to ensure that evaluation methods and tools are used to include rather than exclude vulnerable groups and diverse voices, avoiding biases. For example, remote data collection tools might be more sensitive to bias and exclusion considering its requirements in terms of connectivity, phone access, literacy or language. Detailed risk analysis, stakeholder analysis and close liaison with local partners could help to increase inclusivity. The EM can seek advice from the REO on inclusive and participatory approaches. The evaluation team leader should ensure that considerations around gender and wider equity issues are thoroughly taken into account in the evaluation approach and methods selected, and that these address the diversity of the stakeholders. It should also ensure that data is collected on and from both men and women participants, and gender disaggregated by sex and age. The EM should check that this was done adequately. For more information, see the [TN on integrating gender in WFP evaluations](#) and the [UNEG 2011 Guidance for Integrating Human Rights and Gender Equality](#).

20. **Ethical safeguards:** To be credible, the evaluation methods should be designed and implemented in an ethical way. The implementation of the four ethical principles of Integrity, Accountability, Respect and Beneficence is a shared responsibility among all those engaged in commissioning, hosting, designing, conducting and managing evaluations as well as those subject to evaluation. These have the following specific implications for the evaluation methodology. In terms of accountability, the proposed evaluation approach is expected to address the intended use of the evaluation and outline how stakeholders are engaged throughout the evaluation process. The evaluation methodology should identify and assess ethical risks and mitigating actions. The evaluation team should apply the highest standards of validity and reliability to maximize the accuracy and credibility of evaluative judgements. In terms of integrity, evaluators should be transparent and honest about their methodological or technical knowledge and, during the implementation of the evaluation, information sources for data collection should be selected with due independence and methodological rigor. In terms of respect, the methodology should ensure that relatively powerless, excluded or marginalized groups are given the opportunity and means to be represented in the evaluation. Adequately considering beneficence in the evaluation methods requires to work in a reciprocal manner with informants, not just extracting information, and ensure questions for surveys, focus groups or interviews are value neutral, culturally appropriate and age appropriate. For more information, see the [UNEG 2020 Ethical Guidelines for Evaluation](#).

7. Responsibilities

21. The Evaluation manager holds the main responsibility for the shared decision-making process on evaluation questions, which must take into account the time, data and resources available for the evaluation. He/she is also expected, in the TOR, to make an initial assessment of the data available and provide initial suggestions on the evaluation methods and data collection tools to be considered; at inception phase, s/he shares a document repository with the secondary data with the evaluation team and critically reviews the methodology developed by the evaluation team. In particular, the EM should ensure that the methods selected are appropriate and feasible and seek advice from the REO if needed. At data collection and reporting phases, s/he oversees how the evaluation team applies the methodology through the coordination of the agenda for the fieldwork and the use of the different data collection tools, and how the evaluation team documents it in the report.
22. The evaluation team is expected to fully develop the evaluation approach, methods and detailed data collection tools in the Inception Report to address the evaluation questions raised, making explicit any methodological limitations to the evaluation approach. They are expected to apply during the field work the data collection methods as designed, but with flexibility to accommodate the field reality. At data collection and reporting phases, they should describe the evaluation methodology and limitations in the evaluation report and use the methodological framework developed to analyze the data collected. The details of the evaluation matrix, data collection tools and sample approach should be included in the annexes of the Evaluation Report.

8. Further reading

- Better Evaluation has dedicated pages on a [wide set of methods](#), including on [participatory methods](#), on [analytical designs](#) and on [developmental evaluation](#)
- MEASURE has [dedicated pages](#) on a set of tools for evaluation practitioners.
- [IPDET Handbook Module 8: Data Collection Methods](#)
- [Program evaluation: A Variety of Rigorous Methods Can Help Identify Effective Interventions](#).
General Accounting Office

- [Program evaluation](#): Case Study Evaluations. General Accounting Office
- [USAID evaluation toolkit](#) and [Technical Note on mixed method evaluations](#)
- UNEG Working Group on Evaluation Methods, Compendium of Evaluation Methods Reviewed - Volume 1, December 2020
- [ALNAP \(2016\) Evaluation of Humanitarian Action Guide. ALNAP Guide.](#)
- [American Evaluation Association resources](#), including [Methods for Collecting Information](#)

Academic sources:

- Bamberger, M. (2000) Integrating Quantitative and Qualitative Methods in Development Projects, World Bank Publications
- Patton, Michael Quinn (2010) Developmental evaluation: Applying complexity concepts to enhance innovation and use
- Bamberger, M. et al (2012). Real World Evaluation: Working Under Budget, Time, Data and Political Constraints.
- Greene, J. C. (2007) Mixed Methods in Social Inquiry. Sage Publications
- World Bank, Bamberger et al: Using Mixed Methods in Monitoring and Evaluation: Experiences from International Development.
- Patton, Michael Quinn (2015) Qualitative Research and Evaluation Methods, Sage Publications.
- Patton, M.Q.; McKegg, K.; Wehipeihana, N. (2015) Developmental Evaluation Exemplars: Principles in Practice. Guilford Press.
- World Bank: Case Study Evaluations (L. Morra et al).
- Most Significant Change Technique. A Guide to its Uses. Davies, R. and Dart, J.
- Mayne, J. (2012). Contribution Analysis: Coming of Age? Journal of Evaluation
- Befani, B. (2015). QCA – A Rigorous Qualitative Method for Assessing Impact (How to Note for Coffey).
- Befani, B. & O'Donnell, M. (2016) Choosing Appropriate Evaluation Methods: A Tool for Assessment and Selection. Bond.

For more information, visit our [external](#) and [internal](#) webpages
or contact OEV Cap/Qual Unit at: wfp.decentralizedevaluation@wfp.org