

# Technical Note

## Evaluation Methodology

Version August 2016

### 1. Introduction

1. This TN is intended for use by WFP staff commissioning and/or managing a decentralized evaluation, and/or evaluation teams conducting a decentralized evaluation for WFP. It is intended to **clarify what an evaluation methodology is**, as well as the different potential methodological designs that can be used in WFP decentralized evaluations. The Note **introduces the different methods** which can potentially be used for data collection and analysis in decentralized evaluations; and provides a brief description of what kind of data arises from each method. It also addresses issues such as ethical safeguards in methods, and integrating gender into specific methods.

### 2. What is 'evaluation methodology'?

2. In order to conduct any evaluation, evaluators have to show how they will assess the evaluation criteria and answer the evaluation questions. This entails an **evaluation methodology or a framework which allows evaluators to specify the overall approach to the evaluation (quantitative, qualitative or mixed), what types of data will be collected the methods that will be used to collect the data, and how the collected data will be analyzed**. A good quality evaluation methodology ensures rigor in the evaluation process and produces reliable data enabling the evaluators to make accurate assessments and thus respond sufficiently to the evaluation questions.
3. An evaluation methodology should make explicit:
  - What evaluation criteria and questions will be applied (see [Technical Note evaluation criteria and questions](#))
  - The overall methodological approach to the evaluation (methodology design)
  - The data types and sources to be applied (data types and sources)
  - How data will be gathered (methods including sampling, tools and collection process)
  - How data will be analyzed (analytical methods)
  - How progress will be assessed (see [Technical Note Evaluation matrix](#))
4. The evaluation methodology should be developed **after the evaluation questions are confirmed**, and should be summarized in the **Evaluation Matrix which then serves as a tool to guide application of the design**. The methodology should clearly explain the **rationale** for the evaluation design, data collection methods, sampling and analytical methods to be employed, demonstrating that they are appropriate for the context and purpose of the evaluation and can generate valid findings from a reliable and credible evidence base. The methodology should also state clearly any limitations and the mitigating measures.

### 3. Methodology

5. The methodology should be chosen based on:
  - what **criteria and questions** the evaluation will address;
  - the **evaluation subject** and its context;
  - the **constraints** of the evaluation with regards to time and resources, and/or other factors such as data or ethical considerations
6. The type of evaluation questions affects the methodology to be adopted. There are three main types of evaluation questions as shown in box 1.

#### Box 1: Types of evaluation questions

- **Cause and effect questions:** These questions ask ‘what changes occurred as a result of WFP’s intervention?’ or ‘what would have happened had the intervention not been implemented?’ These can be the most complex questions to answer, as they require an understanding of the situation in the absence of the intervention i.e. counterfactual
- **Descriptive questions:** These questions focus on a particular area and require a descriptive response, for example: ‘Were recipients of WFP’s assistance satisfied with the level of service provided; why did they use or not use WFP services; did women receive different services to their male counterparts?’
- **Normative questions:** These questions assess performance against a specific criterion, for example: ‘has WFP’s specified goal, target or standard, as set out in the Project Document/Logframe, been reached?’

7. In order to respond to these questions, there are several approaches possible as summarized in Table 1 below:

**Table 1: Examples of different approaches to evaluation<sup>1</sup>**

Approach	Uses and types of questions	Features and overall approach	Applications and types of evaluation in WFP
<b>Experimental (e.g. Randomized Control Trial (RCT))</b>	Commonly cause and effect questions such as ‘What is the impact of WFP’s nutrition activity on nutritional indicators in [target population]?’	Two randomly selected groups with a statistically representative sample size  One group receives intervention (treatment group), one does not (control group)	Discrete and specific interventions in a clearly bounded population  Evaluation process must be designed and implemented alongside intervention. This is typically done through impact evaluations
<b>Quasi-experimental (e.g. pretest and posttest for a treated and comparison group)</b>	Cause and effect questions – as above	Does not involve random selection  Uses a natural comparison group, rather than a control  Needs sufficient data to rule out alternative explanations	Can be undertaken with less resources and applied later in the programming cycle; Most WFP evaluation types will include aspects of pre-test and post-test, using baseline information and monitoring data up to the point of evaluation

<sup>1</sup> More discussion of these approaches with examples available [here](#)

<b>Non-experimental e.g. Theory based using Theory of Change, Logframes, or Results Chain</b>	Also cause and effect, as above, but can also be descriptive/normative, such as ‘Did the Operation achieve its intended results’?	Does not compare groups so no comparison group needed Requires understanding of the baseline scenario Builds argument by plausible association between the intervention and observed changes Typically uses mixed-methods	This is a common approach to evaluations in WFP. It requires certain level of baseline and monitoring data, and the presence of documentation that explains the ‘logic’ behind the design of the intervention being evaluated
-----------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

8. These designs have different strengths and weaknesses, as follows:

- **Experimental** – a treatment and control group is randomly selected at the time of the intervention design. The control group does not receive the intervention e.g. a group of non-recipients with the same attributes as a group of WFP target recipients for its activity. The random assignment allows assessment of the impact of the intervention with high levels of validity. However, this approach requires high level of resources, implementation fidelity<sup>2</sup> and specialized skills and evaluation must be designed at the start of the intervention. Finally, the random assignment of control groups may have some ethical issues in the WFP context considering the guiding principle of targeting the most vulnerable in which case the existence of an equally vulnerable group to serve as a control group could bring to question adherence to this principle.
- **Quasi-experimental approaches** involve the use of a ‘**counterfactual**’ group<sup>3</sup> when it is not possible to construct control groups using random assignment. This is called ‘natural comparison group, which is not deliberately pre-selected but is considered to be sufficiently similar to the group receiving the intervention in order for comparisons to be made. This approach requires same level of skills as the experimental in terms of analysis, but does not necessarily have to have been designed at the time of the intervention.
- **Non-experimental approaches**, such as **theory based** evaluations, do not require the use of counterfactual groups. As well as applying to descriptive/normative questions, they can address cause and effect questions by examining the situation prior to the intervention’s implementation, i.e. the baseline scenario, and comparing it with the situation at the time of evaluation. So for example: following five years of Country Programme implementation, what changes are observed that could be connected to WFP’s engagement? Plausible causal links between WFP’s work and the observed changes are then identified. This approach heavily relies on documentation that shows the logic of the design, the theory of intended change or change pathways, then implementation processes and recorded results. As such, where such documentation is weak evaluators have to find ways of reconstructing them, for example through discussions and recalls from key stakeholders.

9. Because of the challenges (methodological, cost, feasibility, and ethical issues) of using counterfactual groups and even natural comparison groups, particularly in humanitarian context and food insecurity, **most WFP evaluations use non-experimental approaches**. However, there is increasing demands for more impact of humanitarian and development interventions. This will require methodological innovativeness to address better cause-and-effect questions, including inter alia better articulation of the design of an intervention theory/theory of change for which plausible

<sup>2</sup> Implementation fidelity is “the extent to which the delivery of an intervention adheres to the protocol or program model as intended by the developers of the intervention”. For further details on implementation fidelity see [here](#).

<sup>3</sup> Counterfactual: The state of the world in the absence of the intervention. 3iE (2012) Impact Evaluation Glossary Version No 7

causal links can be tested through evaluation (see the [Technical Note on Impact Evaluation](#) and [Technical Note on Logic models](#) for more information).

#### 4. Evaluation Methods: Quantitative, Qualitative and Mixed

- Experimental and quasi experimental approaches largely use quantitative data while theory-based approaches use a combination of both. Using a combination of qualitative and quantitative data can improve the quality of an evaluation by ensuring that the limitations of one type of data are balanced by the strengths of another.

#### Box 2: Qualitative and quantitative methods

- Quantitative methods** consist of counts or frequencies, rates or percentages, or other statistics that document the actual existence or absence of problems, behaviors, or occurrences. This data can yield **representative and generalizable** information depending on how it was collected.
- Qualitative methods** can be equally objective and systematic,<sup>4</sup> but the data produced are most often presented in textual form. Qualitative methods often explore ‘how and why’ questions as well as ‘what’ questions.

- Qualitative and quantitative methods have different advantages and disadvantages:<sup>5</sup>

**Table 2: Advantages and disadvantages of quantitative and qualitative data collection methods**

	Advantages	Disadvantages
<b>Quantitative methods</b>	Can include relatively large number of questions and large sample groups and are often more widely accepted as a form of evidence of effectiveness. Is easier to analyze	Lacks the depth and detail and therefore has limited explanatory power. May not provide all the information needed for interpretations of data findings, and large amounts of data may require more sophisticated analysis approaches. Can be harder to design and implement and thus calls for more effort at the planning and preparation phases of the evaluation.
<b>Qualitative methods</b>	Can capture more depth than quantitative methods, e.g. on the “why” and “how” things have happened as a result of WFP’s work. Can clarify quantitative data and help explain it. This makes quantitative data easier to understand, provides more details and nuances, and explains what the intervention has meant to the people involved.	Can be time consuming to capture and analyse. Data resulting can be more challenging to summarize and compare systematically. Generally requires smaller sample sizes. Is harder to analyse.

- Within quantitative and qualitative methods, there are several approaches to gathering the data. Table 3 shows those most relevant to WFP evaluations. Additional sources of information on each method – and on the many other different qualitative and quantitative methods available – can be found in ‘Further reading’, in Box 6.
- A sub-set of qualitative methods are called **participatory**<sup>6</sup>. These participatory methods involve collecting data whilst interacting with people and in which individuals, households, communities

<sup>4</sup> Qunn Patton, M (2015) *Qualitative Research and Evaluation Methods* Sage

<sup>5</sup> Adapted from IPDET (2014) Training Module 8 for evaluators

<sup>6</sup> For further literature on participatory evaluation see [here](#).

and sub-groups are involved directly in the generation of data (and sometimes evaluation questions) for the evaluation. Examples of participatory methods are:

- **social mapping** – can be used to present information on community layout, infrastructure, demography, ethnolinguistic groups, health patterns, wealth and other community issues/resources, climate/natural resource patterns, seasonal calendar etc;
- **Community meetings** – information gathered during meetings of people in the community, such as comments, questions asked, etc.

**Table 3: Tools for each Evaluation Method**

	Tools	Brief description
<b>Quantitative methods</b>	Questionnaires	<ul style="list-style-type: none"> <li>• Questionnaires ask ‘closed’ questions, often focusing on pre- and post- an intervention and therefore relying on recall of the respondents.</li> </ul>
	Statistical research	<ul style="list-style-type: none"> <li>• Gathering information from statistical datasets, which usually already exist (such as national indicators on food security and nutrition in target populations, census, health surveys, etc)</li> </ul>
	Use of quantitative records/ databases	<ul style="list-style-type: none"> <li>• Collecting secondary information from existing databases/in records such as WFP Monitoring or financial data. Requires the development of a structured data collection instrument, and relevant software.</li> </ul>
	Self-reported checklists	<ul style="list-style-type: none"> <li>• Used to gather information about how much time an individual spends doing a particular activity, how frequently a particular activity occurs.</li> <li>• Could be used in a WFP evaluation which includes an element considering staff resources.</li> </ul>
<b>Qualitative methods</b>	Document review	<ul style="list-style-type: none"> <li>• Systematic analysis of different types documents supplied by WFP, Government, donors and partners – should use a clear analytical framework to guide the analysis</li> </ul>
	Key informant or individual interviews	<ul style="list-style-type: none"> <li>• Interviews with individuals, usually applying a structured or semi-structured interview schedule. These are often the main form of primary data collection, involving interviews with WFP’s partners (Government, UN agencies, Co-operating Partner, donors, community leaders). The range of interview targets is informed by the stakeholder analysis.</li> </ul>
	Focus groups	<ul style="list-style-type: none"> <li>• Interviews with groups, usually following a semi-structured format but with scope for open dialogue.</li> <li>• To gain the opinions and views of as many members of the focus group as possible, participatory approaches and tools can be used where appropriate.</li> <li>• Recipients and non-recipients of WFP assistance are often interviewed through focus groups</li> </ul>
	Stories / life histories	<ul style="list-style-type: none"> <li>• Collecting narratives from individuals such as recipients of WFP assistance about their experiences</li> </ul>
<b>Qualitative or</b>	Diaries, journals	<ul style="list-style-type: none"> <li>• Used to gather in-depth information about events in everyday life, and have to be planned in from the start.</li> <li>• These have been applied as a means of monitoring such interventions as support to adherence to anti-retroviral treatment under HIV/AIDS interventions</li> </ul>
	Survey	<ul style="list-style-type: none"> <li>• Can use purely ‘closed’ questions (quantitative) or combine closed-questions with open ended questions (qualitative).</li> <li>• Surveys might be of WFP Co-operating partners or Government counterparts for example.</li> </ul>
<b>Qualitative or</b>	Observation	<ul style="list-style-type: none"> <li>• E.g. of a food distribution. Can be used quantitatively through a coded checklist to record events, behavior etc. such as the extent of under-scooping of food rations, the preferences given to vulnerable groups in the distribution</li> </ul>

<b>quantitative methods (depending on the type of data generated)</b>		<p>process (pregnant/lactating women, elderly and sick). Observation can be obtrusive (participants are aware of what the evaluator is doing), unobtrusive (no-one knows that the evaluator is there and what they are doing).</p> <ul style="list-style-type: none"> <li>• Can also be used qualitatively and/or through a participatory approach, in order to generate an in-depth account of the experience, such as participants' use of vouchers to purchase commodities, or food distributions.</li> </ul>
	Report cards	<ul style="list-style-type: none"> <li>• Used to gather information on how a beneficiary has experienced a particular event or delivery (e.g. a transfer modality).</li> <li>• Can be used to collect quantitative information (e.g. how frequently, how satisfied) or qualitative information (describing the experience).</li> <li>• Usually relies on informants being literate.</li> </ul>

14. There is **no single best way to collect data, and no tool is suitable for all situations**. Evaluation teams conducting evaluations will need to make decisions based upon:

- the evaluation questions;
- where the data reside;
- resources and time available; and
- complexity of the data to be collected.

15. **Mixing methods:** It is now widely recognized that incorporating multiple methods into a single evaluation often results in a stronger, more complete evaluation than conventional evaluation approaches relying on only one method. This trend has led to a rapidly growing interest in mixed-method evaluations among both practitioners and evaluators. **WFP advocates the use of mixed methods in all its evaluations.**

### Box 3: Mixed Methods in Evaluation

A mixed-method evaluation systematically integrates two or more evaluation methods, potentially at every stage of the evaluation process, usually drawing on both quantitative and qualitative data. Mixed-method evaluations may use multiple designs. They also may include different data collection techniques....In short, a mixed-method evaluation involves the systematic integration of different kinds of data, usually drawn from different designs.<sup>7</sup>

## 5. When to consider methods in the evaluation process?

16. For WFP evaluations, how methods are mixed will vary depending on the evaluation questions, the design selected and the context. What is important, however, is that the nature of the mixed method design is clearly signaled and justified in the evaluation methodology outlined in the terms of reference, and further developed during the inception phase and its feasibility critically considered. Methods should be considered by different stakeholders in the evaluation at different times. Table 4 below provides an overview of these.

<sup>7</sup> Source: USAID (2013) [Technical Note: Mixed Methods in Evaluations](#)

**Table 4: Responsibilities for Methodology considerations across the phases**

Phase	Who	Responsibilities
<b>1. Plan</b>	WFP Commissioning Office	To consider which methods might be most appropriate given the evaluation type selected (e.g. for impact evaluation, the presence of a control group or comparison group would be required), the timing
<b>2. Prepare</b>	WFP Evaluation Manager	In the Terms of Reference, specify any particular requirements regarding the methods to be used for the evaluation, and provide a clear outline of the proposed methodology (See <a href="#">Template for Terms of Reference</a> and <a href="#">Quality Checklist for Terms of Reference</a> ). The quality Support mechanism will provide feedback on methodology when reviewing the TOR. The manager should ensure that such feedback is used to strengthen to TOR in terms of methodological clarity. When reviewing proposals (if using consultancy firms to recruit the evaluation team), the EM has to consider whether the methodology proposed by firm is appropriate. If the proposed methodology significantly deviates from that which was outlined in the TOR, the EM must discuss this and come to an agreed way forward before finalizing the contracting of the evaluation team;
<b>3. Inception</b>	Evaluation team	Develop the full evaluation methodology and approach in the form of an Evaluation Matrix (which is also a useful tool for methodology development). Present specific details of data collection and analysis methods in the ‘Evaluation approach and methodology’ section of the Inception Report, and the corresponding data collection tools in the annex see <a href="#">Template for the Inception Report</a> and <a href="#">Quality Checklist for the Inception Report</a>
	WFP Evaluation Manager	Ensure that proposed methods and resources for data collection and analysis are feasible, appropriate, and sufficiently robust and will be applied in a transparent and ethical manner. This is part of the Inception Report quality assurance process. See <a href="#">Quality Checklist for Inception Report</a>
<b>4. Collect Data</b>	Evaluation team	Apply the methods as designed in the inception report, with flexibility when necessary to accommodate reality from the field.  Note gaps in data emerging from particular methods, and implement mitigation strategies on an ongoing basis, documenting them so they can be reflected in the evaluation report
<b>5. Analyze Data and Report</b>	Evaluation team	Analyze the data following the methodology selected and the analytical framework developed during the inception phase In the Evaluation Report, summarize the methodology used, including any amendments or limitations experienced during its application. Include the methodology, including the Evaluation Matrix, data collection and analysis methods, sampling techniques etc., as an annex to the Evaluation Report (see <a href="#">Template for Evaluation Report Template</a> and <a href="#">Quality Checklist for Evaluation Report</a> ).
	WFP evaluation manager	Check that the methods designed during the inception phase were applied for data collection and analysis appropriately in a transparent and ethical manner, and that enough information has been provided in the Evaluation Report to indicate any changes that were made and the rationale  This is part of the Evaluation Report quality assurance process (See <a href="#">Quality Checklist for Evaluation Report</a> )

## 6. Ethical safeguards and considerations in evaluation methods

17. WFP evaluations are often dealing with sensitive subjects, and many involve speaking directly with recipients of WFP assistance who may be highly vulnerable and/or traumatized, as well as partners. All methods used during the evaluation should employ **ethical safeguards** to respect and protect the rights, confidentiality and welfare of the groups and persons involved in the evaluation process. The dignity and diversity of evaluation participants should be respected through use of culturally appropriate data collection methods. Some examples of ethical safeguards and considerations are provided in Table 5 below.

**Table 5: Example of ethical safeguards applied in evaluations**

Ethical safeguard	Description
<b>Informed consent</b>	Participants should be informed of the evaluation purpose and process and be given the chance to consent or not to participate on that basis.
<b>Anonymity</b>	Evaluation participants should be treated as anonymous. A list of key stakeholders interviewed may be provided as an annex to the evaluation report, but this should not include names of recipients of assistance. Within the report, findings should not be attributed to any named individuals. In sensitive cases, organizational affiliations may be listed instead of actual names of persons
<b>Confidentiality</b>	The use for which the data will be made, and the manner in which the findings of the report will be disseminated should be made known to potential participants prior to their giving informed consent.  Interviews with stakeholders should be undertaken without those responsible for the intervention (i.e. WFP and partner staff) being directly present, unless they are the respondents.
<b>Protection and Security considerations</b>	The evaluation team should be aware that their actions, including any explicit or implicit messages transmitted, may have the potential to expose evaluation participants to greater risks, particularly in conflict settings or when working with vulnerable groups. Any risks should be considered and mitigated as far as possible. For example if within the context interviewing women is likely to expose them to undue risks due to cultural practices, the team should find ways to get then women involved without putting them at risk

18. Further information can be found in the [UNEG Ethical Guidelines for Evaluations](#).

## 7. Data types and sources – primary and secondary

19. Data can be gathered from either primary or secondary sources.

20. **Secondary data** is information that is already available, having been collected by someone else for other purposes prior to the start of the evaluation. It is usually in the form of written documents or datasets and can be either internal to WFP or external.

21. Any evaluation requires **some level of existing data**. Part of the preparation process for the WFP Commissioning Office includes gathering existing data for the evaluation team, and identifying any key gaps.

22. A summary of key secondary data sources and major gaps should be provided in the evaluation TOR (see [Template for Evaluation TOR](#): Section 4.3). Evaluation teams should undertake some secondary data collection and analysis prior to finalizing the overall evaluation methodology during the inception phase, as it can provide useful contextual and overview information and thus inform the development of a methodology that is appropriate to the evaluation needs and subject.



23. Secondary data sources can provide useful information regarding the context, for example through: relevant national indicators, policies and strategy documents; WFP and national partner/government plans, strategies and programmes; country studies and other information.
24. Secondary data can also usually provide information about the baseline scenario, for example through baseline studies, baseline indicators in WFP Standard Project Reports, Vulnerability Analysis and Mapping reports and national datasets from prior to the intervention.
25. In order to acquire information on results achieved by the intervention, important secondary data sources include: monitoring data collected during the programme implementation cycle; operational reports (e.g. related to distribution, procurement and logistics); Standard Project Reports and; previous evaluations or reviews reports. The type of monitoring data available will depend on the intervention type and monitoring systems applied.
26. In general, there are numerous relevant types of secondary data sources, depending on the nature and subject of the evaluation. Some key secondary data sources are identified in Box 4 below.

#### **Box 4: Secondary data sources for WFP evaluations**

- Key national and sub-national datasets on food security and nutrition, poverty, education, gender etc.
- Needs assessments, sectoral capacity assessments and risk analyses conducted by WFP and/or other stakeholders, such as comprehensive food security and vulnerability assessments (CFSVA), Joint assessment missions (JAMs), Emergency food security Assessments (EFSAs) etc.
- Relevant WFP policies and guidelines such as Gender, Protection, Food for Assets, Resilience, Cash and Vouchers (see [WFP policy compendium](#))
- Information on the CO structure such as location and number of sub-offices, organigrams, etc.
- Key documents relevant to the transfer modalities and the context of intervention(s), including project documents; assessment reports; SPRs, logframes; evaluations; studies by WFP and other partners; etc.
- Vulnerability Assessment and Mapping (VAM) data
- Monitoring reports including post-distribution reports, on-site monitoring, monitoring plans, monthly monitoring reports etc.
- Data systems e.g. Scope, COMET, LESS, Feedback mechanism system, Retail+ financial service provider datasets
- Maps (activity, food security)
- Procurement and logistics reports (E.g. Pipeline, Projected Needs).
- Notes for Record of Coordination meetings (of clusters, donor groups etc.), humanitarian response plans
- Information related to partners (Field level agreements; Memorandum of Understandings; lists of partners by activity and location).
- Data on interventions/coverage by other actors

27. **Primary data** is collected directly by the evaluation teams for the purposes of answering the evaluation questions. Primary data sources provide more in-depth exploration of results achieved (both intended and unintended) by the intervention and the contextual factors that contributed to them. An advantage of using primary data is that the information has been collected specifically for the purposes of the evaluation – rather than having been gathered for other reasons.
28. Examples of primary data sources typical to WFP decentralized evaluations are provided in Table 6 below.

**Table 6: Primary data sources**

Data source	General description	Typical source in WFP
<b>Data from semi-structured interviews</b>	Data generated from person-to-person interviews following a pre-determined list of questions, with some room for flexibility of discussion.	Key stakeholders from WFP, government, partner organizations, donors, other UN agencies etc.
<b>Expert information</b>	Data gathered from experts related to specific technical areas covered by the evaluation.	Technical experts - usually but not always also stakeholders in the activity/ Operation – such as experts in cash transfer modalities, or in gender
<b>Focus group data</b>	Data gathered from small group interviews (e.g. 6 to 10 participants) for in-depth exploration of specific areas.	Co-operating partner meetings Groups of recipient assistance
<b>On-site observation records</b>	Generated from visits to operational activities and/or project sites to make and record observations.	WFP’s operational activities e.g. food distributions Key project sites
<b>Survey/questionnaire data</b>	Data gathered from a standardized list of questions administered to stakeholder groups	Surveys administered to e.g. groups of Co-operating partners or government partners; Large sample of recipients of assistance

## 8. Sampling

29. It is not usually possible or necessary to interview all stakeholders, review all available documents or visit all relevant field sites of an intervention during the evaluation. Therefore for each data collection method employed, a subset of sites, document and stakeholders must be selected following a clearly developed sampling framework/approach.
30. The mixed methods approach advocated by WFP means that sampling is employed for each method of the evaluation. The overall sampling framework should describe the characteristics of the sample selected for interrogation, how the sample has been selected, how representative it is of the overall population relevant to the evaluation subject, and any limitations that the sampling approach might present. It should be based on specific rationale for selection, which should be made explicit in the evaluation methodology section of the inception report. The rationale can include logistical constraints, for example access issues, but the evaluation team should still ensure that the sample is sufficiently representative of the wider population relevant to the evaluation. For example:
- The rationale for the selection of field sites to visit should aim to show the full and diverse range of the activities WFP is evaluating and their local contexts (rather than simply sites that are close to the capital city, or which provide ‘good examples’)
  - A sample of recipients of assistance should include both men and women, and representatives of relevant marginalized and vulnerable groups where appropriate, as well as a mix of different activities if the intervention being evaluated is an operation with more than one type of activity
31. Samples for qualitative methods are often selected purposively – that is, to make sure that the sample composition contains examples of all the different types of activities/contexts being evaluated. If a sample is not representative, wrong conclusions may be drawn about the larger population. Sample bias should therefore be avoided as far as possible, for example by avoiding the selection process being influenced by those who have vested interests in a positive evaluation outcome. For further guidance, see the [TN on Independence and Impartiality](#). Sample bias can also be reduced by reducing the non-response of selected participants.

32. The sample size should also be considered carefully. While it needs to be realistically based on the resource and time constraints of the evaluation, it must be large enough to ensure that there is sufficient representation of WFP's activities/contexts. For quantitative methods, the sample size may be determined by the need for statistical validity and will depend to some extent on level of homogeneity across the target population. A larger sample size reduces sampling error, which is the likelihood that if another sample of the same size were selected, the results would be different. Once the required sample size to achieve statistical validity is known, the sample is selected following the selected sampling approach i.e. random, systematic or cluster sampling.

## 9. Data analysis and synthesis

33. To answer the evaluation questions, once data has been collected through the methods above, **data has to be analyzed and synthesized**. This is a systematic process of organizing and classifying the information collected, tabulating it, summarizing it and generating findings against the evaluation questions and criteria, from which findings, conclusions and recommendations to respond to the evaluation questions can then be drawn.
34. Data analysis seeks to detect **patterns in the evidence** to help answer the evaluation questions. This can be either through identifying specific individual findings (**analysis**) or combining sources of information to provide a broader understanding (**synthesis**). A **plan for analysis**, including general and specific techniques that will be applied to different data sets for analysis and synthesis, should be clearly articulated in the evaluation methodology in the inception report.
35. Analysis techniques depend on the types of data available. For qualitative data, which is typically in narrative form, an **analytical matrix** might be useful to summarize data from multiple sources and facilitate triangulation and synthesis. Qualitative data analysis software can also be employed. For quantitative data, various **statistical methods** and related software packages can be applied, as appropriate.
36. As part of the analysis, focusing on particular **case studies** (e.g. selected activity types or project sites) can provide a more comprehensive picture of specific aspects of an intervention. The detailed analysis involved in constructing a case study can provide a more in-depth understanding of the specific processes, at a level which may not be possible across the whole of the evaluation subject, for example the Operation. It can also allow for cross examination between different case studies. Case studies should be selected at the methodology design stage, to ensure that sufficient volume and types of data are collected for the chosen cases.
37. Specific methods may be employed for analyzing the cost-efficiency and cost-effectiveness of an intervention. For example:
- **Cost-efficiency** is the analysis of the extent to which the intervention has converted or is expected to convert its resources/inputs (such as funds, expertise, time, etc.) economically into results in order to achieve the maximum possible outputs the minimum possible inputs.<sup>8</sup> This analysis identifies the most efficient among alternatives in achieving intended outputs.
  - **Cost-effectiveness** analysis is a method of comparing the costs and benefits of an intervention, but not necessarily indicated by monetary value.
38. When considering the response to cause and effect questions, causal links between the intervention and observed changes should be established as far as possible. Identifying these causal links ideally allows **attribution** of results to the WFP intervention. However, it is often easier, particular higher up the logic chain (i.e. when considering outcomes and impacts of interventions) to discuss **contribution** rather than attribution. This involves plausible explanation rather than clear 'proof'

<sup>8</sup> [http://siteresources.worldbank.org/EXTGLOREGPARPROG/Resources/grpp\\_sourcebook\\_chap11.pdf](http://siteresources.worldbank.org/EXTGLOREGPARPROG/Resources/grpp_sourcebook_chap11.pdf)

of causal links. Specific techniques in **contribution analysis** have been developed that can be applied in evaluations.<sup>9</sup>

## 10. Triangulation and storing data

39. The use of mixed methods allows for triangulation across different data sources and collection methods. For example, interview responses from project staff can be cross-checked with feedback from recipients of assistance gained during focus groups, and further validated through observations made during site visits and/or secondary data sources.
40. Employing triangulation during the analysis will strengthen the evidence base and confidence in the findings presented, thus improving credibility. It also allows for transparency where findings are supported by limited evidence or there are areas of inconsistency.
41. Prior to the data collection, the likely reliability of different data sources can be estimated and indicated in the Evaluation Matrix. This information, supplemented by knowledge of any factors that may have adversely affected data reliability during the actual data collection activities, should inform the extent to which a data source is relied upon to inform evaluation findings. Information from a less reliable source that does not triangulate with information from other data sources, for example, may be discounted.
42. Depending on the type and quantity of data being collected, specific data management systems may need to be used. As a minimum, a file sharing system is typically established between the WFP Commissioning Office and the evaluation team in order to easily share a library of documents relevant to the evaluation. For evaluations using large sample quantitative methods, the physical and electronic infrastructure that will be used for gathering and handling data should be considered.

## 11. Other relevant reading

43. Methods in evaluation are applied social science research methods; therefore, related research method guides can be a valuable reference.
44. Other relevant reading relevant to methods are listed in Box 6 below.

---

<sup>9</sup> See for example Mayne, J. (2008) Contribution Analysis: An approach to exploring cause and effect, ILAC methodological brief, available at [http://www.cgiar-ilac.org/files/ILAC\\_Brief16\\_Contribution\\_Analysis\\_o.pdf](http://www.cgiar-ilac.org/files/ILAC_Brief16_Contribution_Analysis_o.pdf)

## Box 6: Other relevant reading

- IPDET Handbook Module 8: Data Collection Methods
- Bamberger, Michael (2000) Integrating Quantitative and Qualitative Methods in Development Projects, World Bank Publications
- Greene, Jennifer C. (2007) Mixed Methods in Social Inquiry. Sage Publications
- Journal of Mixed Methods Research. Sage Publications.
- Quinn Patton, M (2015) Qualitative Research and Evaluation Methods Sage
- USAID (2013) Technical Note: Conducting Mixed Method Evaluations
- Burrows, S. & Read, M. (2015). Challenges and Insights from Mixed Methods Impact Evaluations in Protracted Refugee Situations. In Roelen, K. & Camfield, L. (Eds.). *Mixed methods research in poverty and vulnerability: sharing ideas and learning lessons* (pp. 197-230). Houndsmill, Hampshire: Palgrave Macmillan.
- Source: USAID (2013) [Technical Note: Mixed Methods in Evaluations](#)
- For participatory evaluation see [here](#)

**For more information on Decentralised Evaluations visit our webpage**

<http://newgo.wfp.org/how-do-i/do-an-evaluation>

Or contact the DE team at: [wfp.decentralizedevaluation@wfp.org](mailto:wfp.decentralizedevaluation@wfp.org)