How to Plan an Evaluation

Monitoring & Evaluation Guidelines

WFP

United Nations World Food Programme
Office of Evaluation

How to Plan an Evaluation

Overview

Introduction. The purpose of this module is to describe steps, concepts and procedures required for the planning of a country office–managed evaluation.

Why is this Module important?

This module reflects WFP's policy on evaluation, outlining WFP's general evaluation standards and criteria that should be considered in an evaluation. Key steps, procedures and evaluation concepts that provide the backbone for planning a good results-oriented evaluation are described.

The module explains why it is crucial to involve stakeholders in the evaluation process and suggests how this can be done. It also provides guidance on how to select an evaluation team and the timing of key steps in the planning process. In addition, it gives examples of what evaluation objectives, scope, questions and TOR should look like.

What does this Module aim to achieve?

This module has the following objectives:

- To explain when and why evaluations are undertaken at WFP and describe why it is important to involve key stakeholders when clarifying the need for an evaluation.
- To explain the usefulness of preparing a preliminary evaluation timeline and planning checklist, and describe how this can be done.
- To describe the evaluation criteria and standards applied at WFP.
- To describe the link between the logical framework and WFP's evaluation standards and criteria.
- To describe what an evaluation Terms of Reference (TOR) is and what the outline for a good TOR might look like.
- To explain what evaluation scope and objectives are.
- To describe what key evaluation issues might look like and how they are linked to the evaluation objectives.
- To describe the standard data collection methods that can be used in an evaluation.
- To describe key issues that should be considered when selecting the evaluation team.

What should be reviewed before starting?

- What is RBM Oriented M&E
- How to design a Results-Oriented M&E Strategy for Development Programmes
- How to design a Results-Oriented M&E Strategy for EMOPs and PRROs
- How to plan a Baseline Study

Section Titles and Content Headings

- **Clarifying the Need for an Evaluation**
    - Introduction
    - Deciding on the Type of Evaluation
    - Agreeing with Key Stakeholders on the Need for an Evaluation
    - Examples of Stakeholders and Their Potential Involvement in an Evaluation

Clarifying the Need for an Evaluation

Introduction. This section explains when and why evaluations are undertaken in WFP, and it describes why it is important to involve key stakeholders when clarifying the need for an evaluation.

Deciding on the Type of Evaluation

For evaluation, at least 1 of the following 3 exercises must be undertaken during or after the lifetime of an operation lasting longer than 12 months.

Country office self-evaluations should be undertaken for all operations lasting longer than 12 months and should take place prior to the planning of a new phase or at the operations' close.

Country office- or regional bureau-led evaluations (using consultants) should be undertaken for:

- Any operation at any time if a management need arises and if issues cannot be dealt with through self-evaluation; and
- Any operation if the cumulative budget of all phases exceeds US$50 million and if the last evaluation of the operation took place more than 3 years before.

OEDE-managed evaluations should be undertaken for:

- All first-generation Country Programmes;
- Any operation if the cumulative budget of all phases exceeds US$50 million and if the last evaluation of the operation took place more than 3 years before (if such an evaluation is not undertaken by the country office or regional bureau);
- Any operation, thematic or policy evaluation requested by the Executive Board or by senior management.

Agreeing with Key Stakeholders on the Need for an Evaluation

Evaluations should normally be programmed and budgeted during project design and be reflected in the monitoring and evaluation (M&E) workplan. An evaluation may also be programmed on an ad hoc basis by management or at the request of the Executive Board. It is nevertheless important to revisit the rationale for an evaluation, to clarify why it is being undertaken, what issues it will address, and whether its cost will be justified.

Evaluation is first and foremost an important management tool, enabling WFP to review its performance, take decisions, learn from experience and account for its actions, thereby improving the Programme's ability to deliver quality results. Evaluations also support accountability and learning at the country level, by providing stakeholders with the information necessary to assess the performance of WFP and its partners, and to learn and agree on ways of improving the performance of WFP-assisted operations.

Evaluation for accountability purposes looks at past performance to determine the degree of success or failure. Evaluation for learning seeks to improve future performance by determining the reasons for success or failure. WFP evaluations generally contain both elements, with more weight given to one or the other depending on the key purpose and timing of the evaluation. In both cases evaluation is concerned with comparing what has been achieved with what was planned.

In relation to the overall M&E strategy, evaluation is the component that follows up on the baseline study and enables WFP to assess expected and unexpected changes that have occurred as a result of an operation.

When planning and clarifying the need for an evaluation, it is vital that stakeholders are consulted early on, and their role in the evaluation discussed. Experience has shown that such an approach greatly strengthens stakeholder ownership of an evaluation's findings, an important prerequisite for later acceptance of those findings and of the recommendations.

It may be useful to prepare a concept paper as the basis for discussion within WFP and with stakeholders, to help decide whether or not to go forward with an evaluation. A concept paper differs from the eventual terms of reference, as its main purpose is to establish early on the rationale for undertaking an evaluation.

It is also important to be clear on what funds are available for the evaluation, whether additional funds are required, and where those funds might come from. Evaluations are usually funded from direct support costs (DSC). This emphasises the need to programme an evaluation early on and earmark funds for it in the operation approval document.

Country offices can also seek funding and/or team members from interested donors or partners at the country level, or from available grant facilities. The type and scope of an evaluation will be determined partly by the funds available.

Examples of Stakeholders and Their Potential Involvement in an Evaluation

WFP stakeholders can be grouped as follows:
- Full-time and contracted WFP staff;
- Partners at the national level, such as government ministries, United Nations agencies, multilateral and bilateral donors, and NGOs with a national presence;
- Institutional partners engaged in WFP activities; and
- Participants and beneficiaries in WFP activities.

Stakeholder involvement can take many forms – discussing the evaluation concept paper; commenting on the TOR; participating as members of a mission; joining in data collection or analysis, or reviewing and commenting on the evaluation report. It might be useful to establish a small task force of key stakeholders drawn from the above groups. The task force's role would be to assist in preparing the TOR and supporting the evaluation mission.

Examples of Elements That should be covered in a Concept Paper

A concept paper can serve as an early negotiation tool between the country office and stakeholders, setting out expectations and requirements. Suggested elements for inclusion in a concept paper are:
1. Why the evaluation is required (for example, to generate learning for the preparation of the next or another similar operation, or because beneficiary contact monitoring indicates that the operation might not produce the intended results);
2. Who the stakeholders are (e.g. partner agency, government, and donor);
3. The proposed timing of the evaluation;
4. The scope of the evaluation and the key issues to be examined (a brief preliminary outline);
5. The evaluation approach and methods to be used (a brief preliminary outline);
6. Who will undertake the evaluation.

How to Prepare an Evaluation Timeline and Evaluation Plan

Introduction. The purpose of this section is to explain the usefulness of preparing a preliminary evaluation timeline and planning checklist, and illustrate how this can be done.

Example of an Evaluation Timeline and Planning Checklist

Once the need for an evaluation has been agreed upon, it is recommended that the evaluation manager prepares a rough timeline for the implementation of the evaluation. The time it takes to plan and implement an evaluation is often underestimated.

Adequate time is needed to prepare for the evaluation, involve stakeholders in the process and ensure that the results of the evaluation can be used in future operation planning and design. For an evaluation to be credible and produce quality results that are used, good and realistic planning is a must. Below is an example of an Evaluation Timeline and Planning Checklist.

Evaluation Task	When?(insert dates)	How long?	Who?	Done?
TOR drafted	As early as possible. Around 2 months before the evaluation	1 week	Evaluation Manager	
TOR reviewed and stakeholders consulted	See above	2 weeks	Key stakeholders (e.g. Country Office, management, government, partners)	
Team selected and CVs shared among team members	See above	Up to 4 weeks	Evaluation Manager	
Team contracted	As early as possible, no later than 3 weeks before the mission	2–3 weeks	Human Resources Officer	
Background material collected and sent to team members	As soon as team is confirmed	1 week	Evaluation Manager and/or Country Office	
Desk review done	About 1 week before the mission	3–5 days	Team	
Evaluation method outlined	At least 1 week before the briefing of the evaluation team	N/A	Team Leader	
Draft itinerary proposed and agreed upon	Upon agreement of method with team leader	1 week (depending on logistics)	Country Office	
In-country mission	As per agreement	As per TOR	Evaluation team, Country Office	
Start-up stakeholder workshop/meeting prepared and organised	Prepare approximately 1 week ahead	½ day	Evaluation team, Country Office	
Debriefing workshop prepared and organised	Prepare 2–3 days ahead of time during last week of mission	½ – 1 day	Evaluation team and others	

Evaluation Task	When?(insert dates)	How long?	Who?	Done?
Aide-mémoire presented	During debriefing workshops on last few days of mission	Depending on number of meetings	Evaluation team, Country Office, Government	
Draft evaluation report received	Around 3 weeks after departure of team	10 days for team leader; 5 days per team member	Team Leader and team	
Review of and commenting on draft report		2 weeks	Evaluation Manager and relevant key stakeholders	
Comments incorporated and report finalised		1 week	Team Leader	
Recommendation matrix prepared and negotiated	Upon acceptance of final report		Evaluation Manager together with action units	
Report disseminated			Evaluation Manager	

What Evaluation Standards and Criteria are applied at WFP

Introduction. This section describes the general evaluation standards and criteria applied at WFP.

Evaluation Standards and Criteria Applied at WFP

Evaluations must meet the following standards:

- **Evaluations must be useful and used**

  The evaluation must serve the information needs of the intended users. This requires that the needs of all stakeholders be identified and addressed. Reports should clearly describe the operation being evaluated, including its context, and the purposes, procedures and findings of the evaluation. Findings and reports must be disseminated in a timely manner, and implementation of evaluation recommendations must be ensured by the country office.

- **Evaluations must be accurate**

  The evaluation must reveal and convey technically adequate information about the operation, in order to determine its worth or merit. The evaluation report must be evidence based, showing clearly how the evaluation team applied the methods and how the findings were arrived at. Findings must always be triangulated, i.e. supported by several different sources (e.g. key informant, beneficiary and direct observation).

- **Evaluations must seek to reflect the reactions of beneficiaries**

  Evaluation planning must provide for adequate consultations with representative beneficiary groups, with attention given to including the perspectives of males, females, children, and other vulnerable groups as relevant to the operation. Evaluation teams must make use of rapid rural appraisal (RRA) methods whenever possible and should use beneficiary observation and consultation as a key element of their field visits.

- **Evaluations must not be confrontational**

  Evaluations are most effective if done in a constructive manner. Stakeholders should be involved early on and should be allowed to express their information needs. Evaluations must be perceived as helpful and as providing added value, with their key objective being to improve performance.

- **Evaluations must be independent and impartial (unless undertaken as self-evaluations).**

  Evaluators should not have been involved in any stage of the operation being evaluated. The evaluation should be complete and fair in its examination and recording of the strengths and weaknesses of the operation. Differing viewpoints, if they exist, should be presented in the report.

The evaluation criteria that are generally applied by WFP, and that should be reflected in the evaluation objectives, are:

- **Relevance**

  The extent to which the objectives of an operation are consistent with beneficiaries' needs, country needs, organisational priorities, and partners' and donors' policies.

- **Preparation and design**

  The process by which the operation was identified and formulated, and the logic and completeness of the resulting design.

- **Adequacy**

  The adequacy of inputs in relation to the carrying out of the activities.

- **Timeliness**

  The timeliness with which inputs are converted into outputs and outputs are converted into outcomes.

- **Efficiency**

  How cost-efficiently inputs (funds, expertise, time, etc.) are converted into outputs.

- **Effectiveness**

  The extent to which the operation's objectives were achieved, or are expected to be achieved, taking into account their relative importance.

- **Outcome**

  The medium-term results of an operation's outputs.

- **Impact**

  Positive and negative intended or unintended long-term results produced by an operation, either directly or indirectly.

- **Sustainability**

  The continuation of benefits from an operation after major assistance has been completed.

- **Connectedness**

  In the case of EMOP/PRRO only – ensuring that activities of a short-term emergency nature have been carried out in a context that takes longer-term and interconnected problems into account.

- **Coverage and targeting**

  The appropriateness of operation-level targeting of objectives to the local situation, the objectives' compliance with WFP's targeting objectives at the policy level, and the extent to which the planned coverage has been achieved.

- **Partnerships and coordination**

  The appropriateness of the partnerships that have been established with governments, non-governmental organisations (NGOs) and agencies; the effectiveness with which those partnerships have been managed to support the achievement of objectives.

What is the link between the Logical Framework and WFP Evaluation Standards and Criteria

Introduction. This section describes the link between the Logical Framework and WFP's evaluation standards and criteria.

Linking the Logical Framework to WFP Evaluation Standards and Criteria

WFP's evaluation standards and criteria are closely linked to the operation design as outlined in the Logical Framework below.

Logframe level	Information Required	Evaluation Criteria
IMPACT	**Contribution of the operation to wider and long-term objectives.** What have been the long-term changes to the lives of the beneficiaries; can these changes be attributed to the operation or programme? **Have impact indicator targets been achieved/are they likely to be achieved?**	**Impact** - progress towards achievement of long term objectives. *'Making a difference in the long-run.'*
OUTCOME	**Actual achievement of outcome targets compared to the plan.** Were targets achieved - who benefited and how? If targets were not achieved, was this due to poor performance, poorly specified indicators, or problems with operation design; what are beneficiary perceptions of the operation - how do they perceive their lives to have changed? **Have outcome indicator targets been achieved/are they likely to be achieved?**	**Relevance** - Addressing the right problems with the right approach as well as WFP's core mandate responsibilities and national policies. *'Meeting the real needs of beneficiaries.'* **Outcomes** - Extent to which outputs have resulted in the achievement of the outcome or component outcomes. **Effectiveness** - Extent to which an operation has attained its objectives. **Coverage & Targeting** - Have targets been met and the right people benefited at the right time? *'Meeting the right needs.'* **Sustainability** - Prospects for self-reliance and continued utilisation of services after WFP assisted operations have been completed. *'Continuation without outside help.'* **Connectedness** - whether the EMOP/PRRO operation in responding to acute and immediate needs, was taking longer-term needs and problems into account.
	Realisation of assumptions. How did factors outside management control affect achievement of the outcome; did operation design adequately take these into account? Have assumptions been monitored and if so, has this resulted in a change of strategy when indicated?	

Logframe level	Information Required	Evaluation Criteria
OUTPUTS	**Actual achievement of output targets compared to the plan.** Who received food aid and other services; were targets met; could performance have been better achieved through a different approach? **Have output indicator targets been achieved/are they likely to be achieved?**	**Effectiveness and efficiency** -. Testing the quality, quantity and timeliness of outputs and the cost-efficiency with which they were delivered. *'Doing the right things and doing them well.'* **Timeliness** - Testing the timeliness of outputs. **Sustainability** - Prospects for continued delivery of services after WFP assisted operations have been completed. *'Continuation without outside help.'*
	Realisation of assumptions. How did factors outside management control affect achievement of outputs; did operation design adequately take these into account?	
ACTIVITIES	**Actual start-up & completion dates compared with plan.** Beneficiary selection criteria and processes. Was food aid delivered effectively by WFP and its partners; did the organisational arrangements work; were there any delays or time-savings; what effect did any deviations have on the operation?	**Efficiency** - Achievement of an optimum relationship between cost, quality & time. **Timeliness** - Timeliness of activities and processes. **Coordination** & **Partnership** - Taking a joint approach to problem solving and delivery. *'Working well together.'* **Sustainability** - Prospects for continuation of activities after completion. *'Continuation without outside help.'*
	Realisation of assumptions. How did factors outside management control affect completion of activities; did operation design adequately take these into account?	
INPUTS	**Actual input quantities and costs compared to budget.** Were resources provided and utilised according to plan; were inputs provided at least cost and to the desired standards of quality and quantity; if not, how did this affect the operation?	**Adequacy** - Having adequate and timely inputs to carry out activities. *'Securing support and being prepared.'*

Preparing the Evaluation Terms of Reference

Introduction. This section explains what an evaluation Terms of Reference (TOR) is, what you need to keep in mind when you start drafting it and what the outline for a good terms of reference might look like.

Preparing the Evaluation Terms of Reference

The TOR constitutes the plan for conducting an evaluation. A good TOR is an essential ingredient for a successful outcome-oriented evaluation process. Ideally, TORs not only offer clear guidance for the evaluation team but also highlight the expectations of various stakeholders in terms of the expected results of an evaluation. Prior to drafting the TOR, it is important to bear in mind certain principles:

- It is good practice to review all available documentation such as project documents and progress and evaluation reports; these may provide useful insights into the issues the evaluation should address.
- Early on in the planning of an evaluation, the Evaluation Manager must identify the key stakeholders for the evaluation and decide which ones should be consulted on the objectives and key issues to be addressed in the evaluation. Stakeholders should be consulted on the content of the TOR.
- TORs are negotiable. Particularly for emergency and some Protracted Relief and Recovery Operations, experience shows that not all issues are apparent from the beginning but emerge during the evaluation process. Therefore it is sensible to consider undertaking a pre-assessment. For example, the implementing agency could undertake its own rapid self-evaluation, or the Evaluation Manager and Team Leader could liaise with the country office to identify key evaluation issues. Alternatively, time and resources could be allocated to allow the evaluation Team Leader or a team member to make an exploratory field visit prior to the evaluation mission to construct basic data and identify key issues. Findings from the pre-study would then inform the desk study stage of the evaluation, and allow the TOR to be modified accordingly prior to commencement of the main field study phase.
- Ensure that the objectives of the operation to be evaluated are clearly stated. An evaluation will be difficult if it is not clear what the results should be measured against. This is exacerbated in emergency operations, where objectives are likely to be modified as the situation evolves. The following tools are recommended for improving the clarity of operation objectives:
    1. **Logical framework analysis** - which establish a hierarchy of objectives for a particular operation.
    2. **Strategic frameworks** - which set overall objectives for the international community's response, to be adhered to for all emergencies (e.g. United Nations Agency Country Strategic Framework; OCHA Appeal).
    3. **Country or response strategies** - which establishes objectives in the build-up to, or at the onset of, a particular emergency.
- The logical framework plays an important role in operation design, implementation and evaluation. If a logical framework has not been prepared, the Evaluation Manager should include in the TOR a request for the evaluation team to prepare 1.

An Example of an Outline for Terms of Reference

Structure	Contents
Project background	1. A description of the project, operation or theme to be evaluated.
	2. The background for the evaluation, explaining its rationale and purpose.
Project status	3. An update on the current state of implementation.
Purpose and scope of the evaluation	4. A clear statement of evaluation objectives and the scope of the work.
Key issues	5. Key evaluation issues to be covered.
Method	6 .Evaluation method(s) to be used.
	7. Identification of existing reports and performance information.
Team composition	8. Profile and mix of expertise required.
Schedule and logistics	9. Draft itinerary and logistical requirements.
Evaluation report	10. Guidance on and deadlines for expected reporting.

Determining the Scope and Objectives of an Evaluation

Introduction. This section explains what evaluation scope and objectives are. It stresses the need to clarify and prioritise both, and it illustrates how this can be done.

Determining the Scope and Objectives of an Evaluation

Key steps in drafting a terms of reference (TOR) are to clarify the evaluation scope and prioritise objectives. The evaluation scope outlines the broad areas covered (time period, geographic range and range of programme) by an evaluation (and by implication, not covered). The evaluation objectives outline, clearly and concisely, what the evaluation will specifically address within the defined scope. It is important to allow adequate time to discuss and negotiate both the scope and objectives among the various stakeholders involved. However, it is the responsibility of the Evaluation Manager to ensure that the final evaluation objectives are realistic and achievable.

Examples of Evaluation Scope for an Emergency Operation (EMOP)

In determining the scope of the evaluation of an EMOP, the Evaluation Manager may face a number of options:

- Should the evaluation look only at the emergency response or also at actions that were to have taken place prior to the onset of the emergency (e.g. contingency planning)?
- In a prolonged emergency, should the evaluation consider the whole period of the emergency or just the most recent or critical periods?
- Should the evaluation also look at complementarity between the EMOP and other WFP interventions in the same or a neighbouring country?

An Example of Evaluation Scope for a Protracted Relief and Recovery Operation (PRRO)

- This is an evaluation of the WFP portfolio in Angola. As such the evaluation will look at the entire range of activities in the country, including the PRRO, the Special Operations (SOs) and the Quick Action Project (QAP). The main focus and thrust of the evaluation, however, will be on PRRO 6159.0, as it is the largest and most significant project. The SOs and the QAP will be reviewed as they relate and contribute to the overall portfolio, and as they fit into a country-wide WFP and donor strategy.
- The evaluation will also examine the strategic linkages among the United Nations agencies, NGOs, donors implementing partners: first to determine their contribution to meeting the objectives, and second, to shed light on whether the PRRO contributed "to the process of transforming insecure, fragile conditions into durable, stable situations …"

An Example of Evaluation Scope for a Country Programme (CP)

In evaluations of a Country Programme the scope may include:

- Factors affecting the effectiveness of the Country Programme (e.g. review of the country and regional context, the flow of resources made available to WFP and the policies of partner development agencies);
- Compatibility of the Country Programme with the Commitments to Women;
- An examination of the relevance of the activities and the CP as a whole to WFP and the

partner Government's policies and the needs of the beneficiaries; and
- Compatibility of the Country Programme with the Enabling Development policy.
- An examination of the extent to which the Food Aid and Development (FAAD) goals included in the Country Programme have been addressed.

Examples of Evaluation Objectives for an EMOP

- To assess the project impact and draw lessons for the design of future emergency operations.
- To provide accountability to the Executive Board.
- To assess the achievements of the EMOP, and in particular the degree to which the stated objectives are being achieved.
- To assess the evolution in the programming of WFP assistance under this operation, adjusting the targeting mechanisms in order to more adequately address the food insecurity experienced by the displaced and crisis-affected population groups.

An Example of Evaluation Objectives for a PRRO

- To assess the efficiency, relevance, effectiveness, coherence and sustainability of WFP assistance under the PRRO in order to improve implementation of the current operation and assist with planning the next phase.
- To provide accountability to the Executive Board.

Examples of Evaluation Objectives for a CP

- To determine whether the development activities have been designed to make a direct contribution to the objectives of the CP.
- To assess the extent to which the individual WFP activities/projects represent recognised good practice in food aid (including the practices and principles recognised in the Enabling Development policy).
- To provide recommendations that can be used in the development of future Country Strategy Outlines and CPs.
- To provide accountability to the Executive Board.

Examples of Evaluation Objectives for Thematic Evaluations

- To improve the overall management of Special Operations (SOs) by strengthening normative guidance in terms of SO operational policy, processes, and procedures.
- To assess the extent to which the Commitments to Women have been achieved.
- To assess the effectiveness, impact and sustainability of the Commitments to Women.
- To produce recommendations that will help shape WFP's new Gender Strategy 2003–2007.
- To provide accountability to the Executive Board.

Identifying Key Issues and Questions for an Evaluation

Introduction. The purpose of this section is to describe what key evaluation issues and questions might look like and how they are linked to the evaluation objectives.

A Sample of Evaluation Issues and Questions

The most important element of the TOR, and perhaps the most challenging to develop, is the section on the key issues and questions that will be examined in the evaluation. Essentially, this is an exercise in framing evaluation questions against the evaluation objectives identified. The evaluation questions and issues should cover the evaluation objectives in a credible and thorough way.

The table below should be used as a guide to identify the issues and questions that an evaluation may need to raise. These should be limited in number, as too many may result in a diminished outcome and overwhelm the team. In general, be realistic as to what the team can cover, given the available time and resources.

Strategy
● Was the strategy used the most relevant and appropriate to meet the identified needs? ● Were gender issues appropriately reflected in the strategy? ● Were alternative approaches to achieving objectives considered? ● Was the strategy consistent with WFP's core mandate, with WFP's policies and, where relevant, with Country Programme and strategic framework objectives? ● What was the link between food aid and other assistance provided? ● Was the strategy well prepared and did it provide adequate guidance for implementers? For PRROs: Did the strategy accurately consider the opportunities to introduce recovery activities? ● Was an exit strategy developed? ● Were other agencies included in the analysis of the problem and the drawing-up of the response plans? ● Did the preparation process adequately involve beneficiaries (both men and women)? ● Were constraints to implementation adequately assessed from the onset? ● Has the strategy been periodically reviewed or modified over the life of the operation in order to maintain its relevance to changing circumstances?
Design
● Was the design of the operation logical and coherent? Did it provide a good "road map" for implementation and a sound basis for review and evaluation? ● Were objectives set at the right level (i.e. in compliance with logframe definitions) and were sound means-to-ends linkages made between inputs, activities, outputs, outcome(s) and the impact? ● Was the operation design technically feasible? ● Were gender concerns appropriately reflected and in line with the Commitments to Women? ● Were assumptions reasonable when they were specified, and were contingency plans made for known risks? ● Which external factors were not taken into account during design? ● Has the operational context of the operation changed since its design, and if so has the operation adapted effectively?
Achievement of objectives

● Are the objectives appropriate and realistic? ● To what extent have planned outputs been achieved in terms of quantity, quality, equity and timeliness? ● For EMOPs and PRROs: have the targeted beneficiaries received the planned rations? Have the right beneficiaries been targeted? Did the beneficiaries use the food aid as intended? ● For PRROs and development activities: have the beneficiaries maintained or improved their livelihoods and what assets have been created/maintained, who have they benefited and how? ● What were the most successful recovery activities? What were the least successful? How effectively was a transition made from relief to recovery activities? To what degree has food aid contributed to promoting resettlement and food self-sufficiency (where applicable)? ● Have outputs reached the poorest? Women? Vulnerable groups? To what extent?· What share of women and vulnerable groups have benefitted directly from FFW outputs? ● Were resources effectively and efficiently used? Do the results justify the costs? ● What evidence is there that the operation's stated purpose(s) and goal were achieved? ● To what extent were WFP's core mandate responsibilities met? ● If at mid-term, what additional measures are needed to improve the chances of achieving the purpose?
Factors in the effectiveness and efficiency of EMOPs and PRROs
Registration/verification ● Were data collected for target populations/groups reliable and disaggregated, and how well were the data maintained? **Assessments** ● How was the initial situation assessed? How has the beneficiaries' food insecurity, vulnerability and nutrition status been assessed? Were the roles of men and women adequately analysed? Have regular reassessments been conducted? Have host communities' needs been assessed? **Final distribution** ● Was the delivery system efficient and equitable? Was it transparent? Was a complaints procedure in place? How well were failures in the delivery system addressed? ● Did women play lead roles on local decision-making committees for food management? If so, to what extent, and has this resulted in a change in acceptance of women in lead decision-making roles?
Targeting
● How have beneficiary groups/areas been identified? ● Were the special needs of certain groups/areas considered? ● Have the needs of the host community been addressed? ● Were targeting objectives appropriate? ● Were arrangements made to review and update the targeting strategy? ● Were male and female beneficiaries involved in the negotiation of beneficiary status? ● What was the proportion of male and female beneficiaries participating in food for work (FFW)?
Appropriateness of rations
● Was the food ration and basket adequate and acceptable (quantitatively and qualitatively) vis-a-vis the livelihood and coping strategies of both men and women? ● Were the nutritional objectives realistic, and to what extent were they achieved? ● What has been the nutritional effect/outcome of WFP assistance on beneficiaries? ● Have there been ration reductions or phasing-out of food assistance; if so, on what basis? ● Was the food aid culturally acceptable and appropriate?
Factors in the effectiveness and efficiency of all operations

Implementation

- Were there any significant delays in approval; what were the causes of this?
- Were the implementation schedule and management arrangements realistic (this should include financial management and budgeting systems)?
- Were plans followed? If not, why not?
- Did objectives change during implementation? Why? On the basis of what analysis or what events?
- Have there been delays in implementation? How were these dealt with? What was the impact of such delays?
- How well were risks and problems managed?

Monitoring and reporting

- Were baseline data collected and were appropriate indicators identified at the outset for measuring progress and results?
- Did the monitoring systems work, including the extent to which gender considerations were built into monitoring arrangements? Were gender-disaggregated data collected in accordance with WFP policy?
- What are the type and frequency of reporting for the operation, including periodic participatory appraisals? Were reports submitted on time? Were they complete? What problems were experienced? Is the information analysed and used to make decisions regarding the management of the operation?
- For PRROs: have criteria been established to signal when to shift activities from relief to recovery, and from recovery to development? If so, are these being applied effectively?

Coordination, partnerships and beneficiary participation

- Were objectives and activities compatible with the policies/programmes of the relevant government(s)? How supportive are WFP's counterparts?
- What mechanisms were put in place for coordination with the Government, donors, United Nations agencies, NGOs, etc. and how effective were they? Were sectorial responsibilities successfully divided among partners?
- Did implementing partners fulfil their contractual obligations? How was their capacity assessed? What training of partners has taken place? Has there been a trend towards or away from using local implementing partners? What capacity has been left behind?
- In what way have the beneficiaries participated in the design, implementation and monitoring of the operation? Were the arrangements adequate? What was the impact of their participation or non-participation?
- Was beneficiary involvement empowering men and women?

Programmatic linkages

- Are the objectives and activities of the operation/programme compatible with and complementary to those of the other operations currently being implemented by the country office?

Management, human resources and training

- Did the country office, sub-offices and specialised units fulfil their roles effectively and efficiently?
- Did staff have the experience and expertise to carry out the activities envisaged?
- Are there staff or skill shortages? What has been done about these?
- Have correct financial and administrative procedures been followed?
- Is there appropriate gender balance in the operation staff (both WFP and implementing partners)? Have staff been trained in gender issues and are they applying this training to their work?

Pipeline, commodity control and logistics

- Has there been efficient pipeline management?
- Commodity sourcing and delivery: what difficulties and losses were experienced; how were these resolved?
- Was planning for logistics, transport, storage, staffing and auditing adequate? What have been the major challenges to the smooth functioning of the operation/programme?
- For EMOP/PRRO: What has been the overall scale of the war economyand the relative scale of contribution to it by diverted or taxed humanitarian assistance? What have been the steps taken by WFP and partners to minimise the level of food diversion or taxation through the selection of particular delivery channels, supervised distributions and end-use monitoring?

Security

- Have adequate and appropriate measures been introduced and adhered to, in order to minimise the risk to WFP staff and implementing partners involved in the implementation of the operation/programme?
- Have measures been taken to provide protection to the men, women and children participating in the operation?
- Are there significant security challenges to the smooth functioning of the operation/programme? Was the safety of the beneficiaries a high priority?

Budget and resources

- Operation costs: What was the planned vs. actual expenditure for the operation as a whole, and under each main cost heading?
- Was the budget appropriate in relation to its objectives/activities, and what factors affected individual budget items, particularly direct support cost (DSC)?
- ITSH: How were internal transport, storage and handling financed? Were the calculated rates adequate? What revisions were required?
- To what extent have the resourcing requirements for the operation been met and how has the country office managed shortfalls? How predictably and regularly have resources been supplied to the operation/programme?
- Monetization: Was it justified? Were the prices fair?
- For PRROs: Has transformation to a PRRO resulted in longer-term (more than one year) financial commitments to the operation?
- How successfully has the operation/programme resourced its non-food inputs (NFIs) and what, if any, have been the constraints?

Standards and quality

- What systems were put in place for assuring programme quality, including setting appropriate technical standards?
- Have relevant international standards and code of conduct been addressed?

Outcome/impact and sustainability/connectedness
● What difference did the operation make to the lives of the beneficiaries (nutritional status, livelihood improvements, etc.)? Who benefited exactly? What would have happened if no assistance had been given? ● How do beneficiaries and other stakeholders value the operation? ● What impact did the operation have on household food security and self-reliance? ● EMOPs and PRROs: What was the impact on host populations? ● What unexpected outcomes (positive or negative) have there been? ● Have environmental concerns been adequately addressed within the operation/programme? ● What intended or unintended impacts have occurred on the environment as a result of the operation? ● What impact have campsites had on the environment? What effect have rations and cooking requirements had? ● What was done to ameliorate the operation's environmental impact and restore the natural habitat? ● How long are these benefits likely to last? What are the prospects for the sustainability of activities and outputs? What factors are likely to undermine sustainability of benefits? Are the sustainable benefits likely to exceed the costs? ● What costs will have to be borne by government, implementing partners, beneficiaries or other stakeholders in order for the benefits to continue? ● Has food aid been used as leverage to obtain complementary national and international resources and recognition to improve the condition of women?
Commitments to Women
● To what extent were WFP's Commitments to Women met in terms of (i) food access by women; (ii) access to power structures and decision-making; and (iii) access to resources, employment, markets and trade? ● What changes are required in a future phase to ensure better compatibility with the Commitments? ● What gender training has been carried out, for whom, and what difference, if at all, has such training made? ● Was adequate effort been made to mainstream as well as advocate gender considerations?

What are the Standard Methods used in an Evaluation

Introduction. This section describes the data collection methods that can be applied in an evaluation. It states the need to relate the planned data collection methods to the evaluation questions, and to the data collection methods applied in the baseline study, if any.

What Standard Data Collection Methods are Used in Evaluations

The terms of reference (TOR) should contain a brief outline of the proposed methods for an evaluation. The data collection methods selected should ensure that the evaluation questions can be fully and credibly answered. If a baseline study has been conducted and the evaluation is a follow-up study to that baseline, the data collection methods applied in the baseline study should be adopted in the evaluation. This will facilitate comparability and enable WFP to assess expected and unexpected changes that have occurred as a result of an operation.

There are essentially 5 ways of obtaining evaluation information:

1. Collecting, tabulating and reviewing existing data;
2. Questioning people through interviews and focus groups;
3. Facilitating critical performance and results reflection by implementers and beneficiaries;
4. Conducting surveys; and
5. Observing people, their interactions and environment through field visits.

Interviews with a sample of beneficiaries or with the affected population can be 1 of the richest sources of information for the evaluation, and are an absolute must. Planning for beneficiary and stakeholder involvement requires, at minimum:

- Scheduling sufficient time and resources;
- Ensuring that the evaluation team has the required skills and experience;
- Facilitating the meaningful involvement of stakeholders and beneficiaries through individual and group interviews, and through the use of rapid rural appraisal (RRA) and participatory rural appraisal (PRA) techniques; and
- Facilitating critical performance and results reflection by implementers and beneficiaries (the evaluator turns into a facilitator).

Examples of Standard Data Collection Methods for an Evaluation

Examples of standard data collection methods are elaborated in the table below, which includes comments on the applicability of each method.

Document review	**File reviews** Essential starting point for all evaluations. Involves examination of available reports, files and other project documents. Good for familiarization with the operation to be evaluated, the development of a historical perspective, and the identification of an initial list of evaluation questions.
	Literature search Economic and efficient way of obtaining information. Difficult to assess validity and reliability of secondary data.
	Tabulating data Registration of quantifiable or classifiable data by means of analytical instruments. Analysis of performance indicators tracked by operation management. Precise, reliable and often requiring few resources. Registers only facts, not explanations.
Consultation	**Key informant interviews** Flexible, in-depth approach. Easy to implement. Risk of biased presentation/interpretation from informants/interviewer.
	Timeline Useful tool for initial meetings with key informants and focus groups. A timeline of important events is created, to help reconstruct events and explore the perceptions of different stakeholders regarding the sequence and importance of those events.
	Focus Group Interviews For analysis of specific, complex problems, to identify attitudes and priorities in smaller groups. Reasonable in terms of cost, and efficient. Stimulates the generation of new ideas. Risk of domination by individuals and of bias of moderator.
	Group interviews Low-cost, efficient. Direct contact with those affected. Susceptible to manipulation and less suitable for sensitive issues.
	Other RRA/PRA techniques Visual techniques, such as mapping, ranking and scoring, and verbal techniques, such as transect walks, provide a means of generating a local analysis, and involving beneficiaries and stakeholders directly in the evaluation process.
	Self-evaluation of performance and results obtained The evaluator becomes a facilitator, facilitating a critical self-assessment of implementers and beneficiaries. Less extractive, with stakeholders potentially becoming owners of the evaluation process.

Survey-based techniques	**Formal survey** Oral interviews or written questionnaires in a representative sample of respondents. Data collection is demanding but often produces reliable information.
	Informal survey Involves quantitative surveys of small samples. Reasonable and rapid. Risk of sampling errors/biases. Less suited for generalization.
	Case studies In-depth review of one or a small number of selected cases. Well suited for understanding processes and for formulating hypotheses to be tested later.
Observation	**Direct observation** Involves inspection, field visits, and observation, for understanding processes, infrastructure/services and their utilisation. Dependent on observer's understanding and interpretation.
	Observation In-depth observations over an extended period of time, participatory or non-participatory. Well suited for understanding processes but with limited potential for generalisation.

How to select the Evaluation Team

Introduction. This section provides guidance on how to select an evaluation team.

How to select the Evaluation Team

When using external consultants in an evaluation, it is advisable to start the recruitment process well in advance of the date of the evaluation mission, in order to ensure the availability of team members and allow time for contracting. If the dates for the evaluation are already set, the search for consultants can be done even while negotiating and preparing the terms of reference.

At times, a donor or partner agency may request that 1 of their own staff join the evaluation mission. This is possible as long as the person proposed fits the criteria outlined below, and functions as a full team member under the guidance of the team leader. WFP should reserve the right to reject candidates if they do not meet the selection criteria. Where larger teams are required (for example, to conduct collaborative or system-wide evaluations) the team leader must possess strong team management skills.

Evaluations of Emergency and Protracted Relief and Recovery Operations will require that considerable time be spent interviewing beneficiaries, in particular, women and children. In cases where beneficiaries have experienced serious psychosocial trauma, it is important that the evaluation team include adequate gender balance and members with appropriate expertise to undertake the interviews.

Examples of Selection Criteria for an Evaluation Team

Evaluation team members should usually fulfil the following requirements:

Team members	• Professional competence and solid technical experience in areas such as food security, nutrition, gender, school feeding, emergency operations and food for work. The expertise required will depend on the key issues to be looked at by the evaluation • Knowledge of WFP and its operations (at least 1 team member, preferably the mission leader, should have WFP experience) • Prior evaluation experience and familiarity with standard evaluation methods • Participatory rural appraisal (PRA) and rapid rural appraisals (RRA) expertise • Impartiality (e.g. not belonging to 1 of the implementing agencies, not having participated as a consultant at any stages of the operation, not being involved in the country in any other or related official capacity) • Preferably prior experience working in the country (at least 1 consultant must be a national of the beneficiary country) Language skills (English, French, Spanish and local languages)
Team leader (in addition to the above)	• Proven team-leading skills • Proven excellent writing skills in English, French, or Spanish • Proven track record as an evaluator and team leader • Credible
Other criteria	• Mixed professional international and national expertise • Gender (appropriate consideration of qualified female consultants) • The team must be balanced in terms of gender, nationality and skills

Module Summary

What has been covered in this module?

Key steps, procedures and evaluation concepts that provide the basis for planning a good results- oriented evaluation were covered in this module. It explained why it is crucial to involve stakeholders in the evaluation process and it suggested how this can be done. Terms such as evaluation scope, objectives and questions were explained, and their link and importance in the planning of an evaluation was clarified.

What additional resources are available?

For further information the following modules and resources might be useful:

- How to plan a Baseline Study
- How to manage an Evaluation and disseminate its Results
- Choosing Methods and Tools for Data Collection
- Going to the Field to collect Monitoring and Evaluation Data
- How to consolidate, process and analyse Qualitative and Quantitative Data
- Reporting on M&E Data and Information for EMOPs and PRROs
- Reporting on M&E Data and Information for Development Programmes
- How to plan and undertake a Self-evaluation
- OEDE's generic TOR for PRRO Evaluations

How to Plan an Evaluation

United Nations
World Food Programme
Office of Evaluation and Monitoring

Via Cesare Giulio Viola, 68/70 - 00148
Rome, Italy
Web Site: www.wfp.org
E-mail: wfpinfo@wfp.org
Tel: +39 06 65131