WFP EVALUATION

World Food Programme

SAVING LIVES
CHANGING LIVES

Guidance for process and content

Decentralized Evaluation Quality Assurance System

December 2024

Contents

Foreword

The Decentralized Evaluation Quality Assurance System (DEQAS) is one of the building blocks for implementation of the 2022 WFP Evaluation Policy. As such, it is the evaluation function's primary means of safeguarding international evaluation principles of:

- **INDEPENDENCE**: by setting standards that increase the impartiality of the evaluation process and reporting on findings.
- **CREDIBILITY**: by setting standards that ensure evaluations are evidence-based and follow transparent and systematic processes.
- **UTILITY**: by building milestones into evaluation processes for timeliness and reporting standards to ensure accessibility.

The DEQAS guides all decentralized evaluations commissioned by WFP and undertaken by independent evaluation teams. The DEQAS is a system to support and guide country offices, regional bureaux and headquarters divisions (other than the Office of Evaluation) when planning, commissioning and managing decentralized evaluations. The DEQAS is designed for the WFP evaluation cadre and evaluation teams, covering all stages of the evaluation cycle. It does not replace the range of evaluation literature available.

The DEQAS builds on the norms and standards of the United Nations Evaluation Group (UNEG); the Organisation for Economic Development's Development Assistance Committee (OECD-DAC) Evaluation Network; related tools from the Active Learning Network for Accountability and Performance (ALNAP); and wider evaluation literature and community of practice.

The DEQAS pack consists of:

i. **Process and contents guide**
ii. **Quality checklists**
iii. **Templates**
iv. **Technical and guidance notes**
v. **Mini-guide for country directors/deputy country directors**
vi. **Other reference materials**

The first version of the DEQAS was piloted in 2016. Since then, it has been periodically and systematically updated in line with the WFP evaluation function's evolving needs, international best practice and feedback from regional evaluation units (REU)s and evaluation managers (EMs) to ensure that the final material serves its intended purpose. In 2020, the Office of Evaluation (OEV) reviewed its entire Evaluation Quality Assurance System (EQAS) across all types of evaluations to ensure closer alignment where required and to reflect recent UNEG guidance, such as the 2020 Ethical Guidelines for Evaluation.

Since 2020, the DEQAS has been reviewed to define changes related to multi-country, regional and headquarters-led decentralized evaluations; to reflect new guidance in terms of diversity and inclusion as well as data protection; and to incorporate good practices in each evaluation phase.

Further updates and new material will continue to be added as needed, to ensure the DEQAS continues to reflect emergent best practice and changes in WFP evaluation systems and processes.

Anne-Claire Luzot
WFP Director of Evaluation

At intervention design stage	Up to 9 weeks	Up to 7 weeks	Up to 3 weeks	Up to 11 weeks	Up to 4 weeks
1 PLANNING	**2 PREPARATION**	**3 INCEPTION**	**4 DATA COLLECTION**	**5 REPORTING & ANALYSIS**	**6 DISSEMINATION**

USE

PRODUCTS

1	2	3	4	5	6
MONITORING REVIEW & EVALUATION PLAN	TERMS OF REFERENCE COMMS PLAN DOCUMENT LIBRARY	INCEPTION REPORT	DE-BRIEFING PRESENTATION	EVALUATION REPORT	LEARNING PRODUCTS

KEY ACTIONS

1 Planning
- Identify eval. type, timing and scope
- Plan + budget in CPB
- Appoint EM
- Set up EC
- Identify ERG members

2 Preparation
- Enroll in EvalPro
- Prepare Stakeholder Matrix and consult stakeholders
- Establish ERG
- Identify eval. objectives & questions
- Draft ToR
- Prepare document library
- Develop Communication and KM plan
- Quality assure draft ToR
- Approve final ToR
- Hire Eval. team and finalize budget
- Prepare Contract/ Purchase Order & collect signed Pledge of Ethics + confidentiality agreement

3 Inception
- Conduct team orientation
- Undertake desk review
- Hold inception meetings with stakeholders
- Draft IR
- Quality assure draft IR
- Approve IR
- Circulate IR

4 Data Collection
- Prepare fieldwork/ schedule field visits
- Provide logistical support
- Conduct field work and preliminary analysis
- Hold end of mission debriefing

5 Reporting & Analysis
- Draft ER
- Quality assure draft ER
- Approve ER

6 Dissemination
- Develop MR
- Review MR
- Approve MR first level
- Approve MR final
- Publish ER + MR
- Disseminate eval. evidence and results
- Submit final ER to PHQA
- Develop and disseminate learning products
- Track implementation of follow-up actions to recommendations

TOOLS

1 Planning

Technical Notes
-Decentralized Eval. Types
-Decision Guide for Impact eval.
-Joint evaluation
-Eval. Reference Group
-Eval. Committee
-Eval. Planning & Budgeting
-Engaging with Donors

Templates
-Budget & Timeline Form
-Eval. manager ToR

2 Preparation

Technical Notes
-Options for contracting eval. teams
-Eval. Criteria and Questions
-Eval. Reference Group
-Eval. Committee
-Contingency Eval. Fund (CEF)

Templates
-ToR
-Summary ToR
-Comments Matrix
-Communication & KM Plan
-CEF Application Form
-Offer DE Services

Quality Checklists
-ToR

Tools
-Assessment of Eval. Proposals
Info brief on cancelling DEs

3 Inception

Technical Notes
-Eval. matrix
-Logic models
-Eval. Approaches, methods & tools
-Formatting Guidelines

Info Brief
-Developmental eval.

Templates
-Inception Report
-Comments matrix

Quality Checklists
-Inception Report

4 Data Collection

Templates
-End of Field Work Debriefing

Other reference docs
-WFP Guide to Personal Data Protection

Guidance Notes
-Remote data protection

5 Reporting & Analysis

Technical Notes
-Quality of Eval. Recommendations

Templates
-Eval. Report
-Comments matrix
-Baseline Report
-Eval. Brief

Quality Checklists
-Eval. Report
-Baseline Report

6 Dissemination

Technical Notes
-Management response to DE
-Communication & KM Plan

Templates
-Management response
-Eval. Brief
-End of Eval. Lessons Learned Calls Info Brief and Templates
-Tag eval. recs against themes

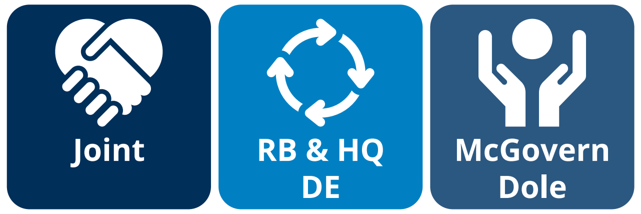
Cross-Cutting
Gender TN, Checklist, Quick Guide | Eval. Principles, Norms and Standards | Planning and Conducting Eval. during COVID-19 | Stakeholder Analysis | Methodology | Data Protection info brief | Disability inclusion | Glossary of Terms | MIS Guidelines

Cross-Cutting UNEG Documents
Ethical Guidelines for Eval. | Guidance for Integrating Human Rights and Gender Equality | Disability inclusion

Introduction

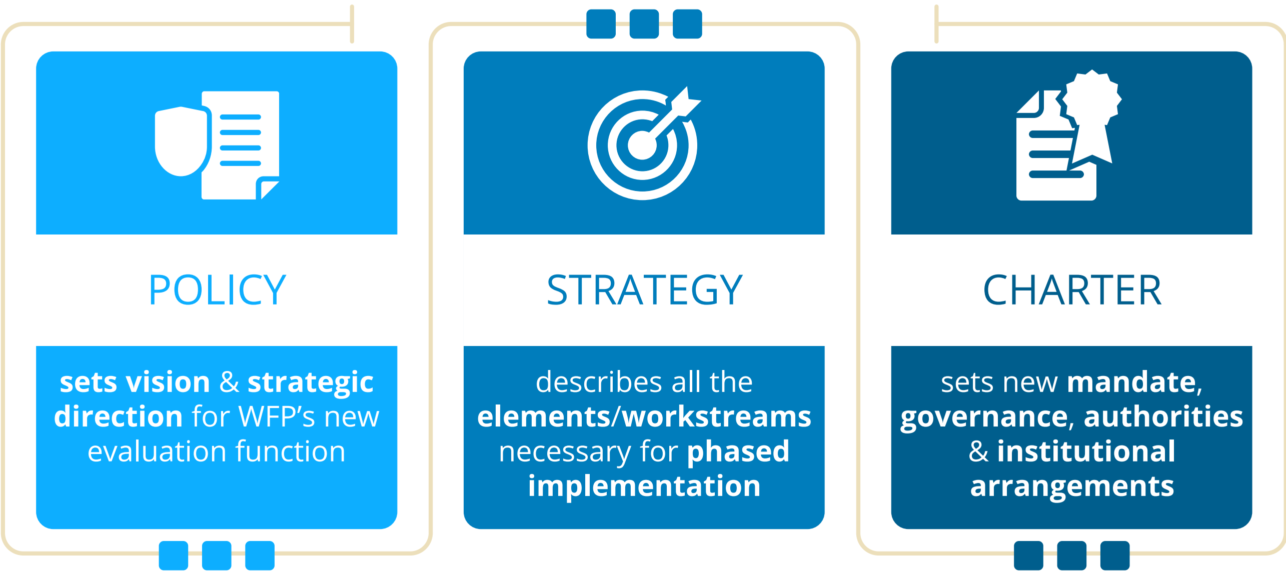
The DEQAS Process Guide is intended for WFP staff managing decentralized evaluations (DEs) commissioned by country offices (COs), regional bureaux (RBs) and headquarters divisions/units commissioning other than the Office of Evaluation (OEV). The guide is intended for three categories of users: evaluation managers (EMs) of decentralized evaluations at country offices, regional bureaux and headquarters (HQ) divisions; regional evaluation units (REUs) for guiding and supporting country office evaluation managers, as well as chairs of evaluation committees and evaluation teams; and the Office of Evaluation for guiding and supporting managers of decentralized evaluations commissioned by headquarters.

Because the majority of the decentralized evaluations are commissioned by country offices, each section starts by explaining the processes for a country office-commissioned decentralized evaluation, then signals what would be unique/different for a decentralized evaluation commissioned by a regional bureau or headquarters, or a multi-country decentralized evaluation (MCDE). There is also a signal on what would be different/unique for joint decentralized evaluations as well as the decentralized evaluations of McGovern Dole (MGD) projects, which have some unique requirements. These three elements are signposted as shown below:



The guide is structured around the six phases of an evaluation process, from the planning phase through to the follow-up and dissemination phase. The guide sets out the specific steps of each phase with links to relevant materials. To get quickly oriented to the guidance and decentralized evaluation process, especially for new decentralized evaluation managers and/or those coming back to manage a decentralized evaluation after a while, it is recommended to review the process map on page 6 in detail (print a copy for reference) and review the tips at the start of each phase to learn what can go wrong and some good practices. These have been gathered from feedback of those who have several years' experience managing and/or supporting decentralized evaluations.

Figure 1: WFP Evaluation framework



1. BACKGROUND

The WFP evaluation function is framed by three foundational pillars (see Figure 1):

- The Evaluation Policy 2022, approved by the Executive Board, sets the vision, strategic direction and model for the WFP evaluation function – that WFP contribution to achieving zero hunger is supported by evaluative thinking, behaviour and systems.
- The Evaluation Charter 2023, issued by the Executive Director, confirms the mandate and governance of the evaluation function, and establishes the necessary staff authorities, roles and institutional arrangements to operationalize the policy.
- The Corporate Evaluation Strategy (CES) 2022 sets out a phased implementation plan, comprising all the elements and activities required for building a combined centralized, impact and demand-led decentralized evaluation function that meets United Nations evaluation norms and standards, and achieves the vision set out in the Evaluation Policy.
- Each of the six regional bureaux articulates how the Corporate Evaluation Strategy will be implemented across the region in the Regional Evaluation Strategies (RESs) which are aligned with the Corporate Evaluation Strategy.

These documents clearly outline WFP commitment to demand-led decentralized evaluations, complementing centralized evaluations and impact evaluations in line with the organization's commitment to enhancing evidence and accountability for results. All country offices are expected to commission at least one decentralized evaluation during each country strategic plan (CSP) cycle but can and do commission more than one to meet their evidence needs. The regional evaluation units consolidate all planned evaluations across the region in the regional evaluation plans (REPs) on a yearly basis. Endorsed by the regional evaluation committees (RECs), the regional evaluation plans set out evaluation priorities for each region. Headquarters divisions may also commission decentralized evaluations as needed. In addition, the Office of Evaluation (OEV) commissions policy, strategic, CSP, corporate emergency response, impact and synthesis evaluations.

To ensure their independence, all WFP evaluations are conducted by independent consultants. While some organizations refer to the term "self-evaluations" or "internal evaluations" when they are conducted by their own staff, WFP refers to these exercises as "reviews" and does not consider them evaluations.

The evaluation principles of independence, impartiality, credibility and utility underpin WFP evaluation framing documents. The Technical Note on Principles, Norms and Standards provides information to help apply these principles in the design, conduct and management of high-quality evaluations. See also Table 1 below and step 1.4: Clarify provisions for impartiality and independence.

Table 1: International evaluation principles

Independence and impartiality	As stated in the UNEG Ethical Guidelines, "independence and impartiality prevent bias and are therefore both essential for the credibility of an evaluation. The main difference between the two is that independence relates to external pressure or influence on those who organize or do evaluations, while impartiality is an attitude towards the evaluated subject. In other words, independence means that the evaluator should be able to work freely and without outside interference, while impartiality means that the evaluator should not be biased with respect to what is being evaluated." This implies that the evaluators must not have been (or expect to be in the near future) directly responsible for the policy setting, design or management of the evaluation subject.
Credibility	Credibility is the extent to which evaluation findings and conclusions are fair, impartial and complete. Credibility is determined by the independence, impartiality, transparency, methodological appropriateness and rigour applied in evaluations.
Utility	Utility is the extent to which evaluations are useful to decision makers and stakeholders, informing policies, strategies and programmes, and in meeting accountability requirements. WFP is committed to enhancing utility by planning and conducting evaluations with a clear intent to use their results; undertaking evaluations in a timely manner to inform decision making processes; and ensuring the accessibility of evaluation results by making reports publicly available.

2. OVERVIEW OF DECENTRALIZED EVALUATION SUPPORT MECHANISMS AND GUIDANCE

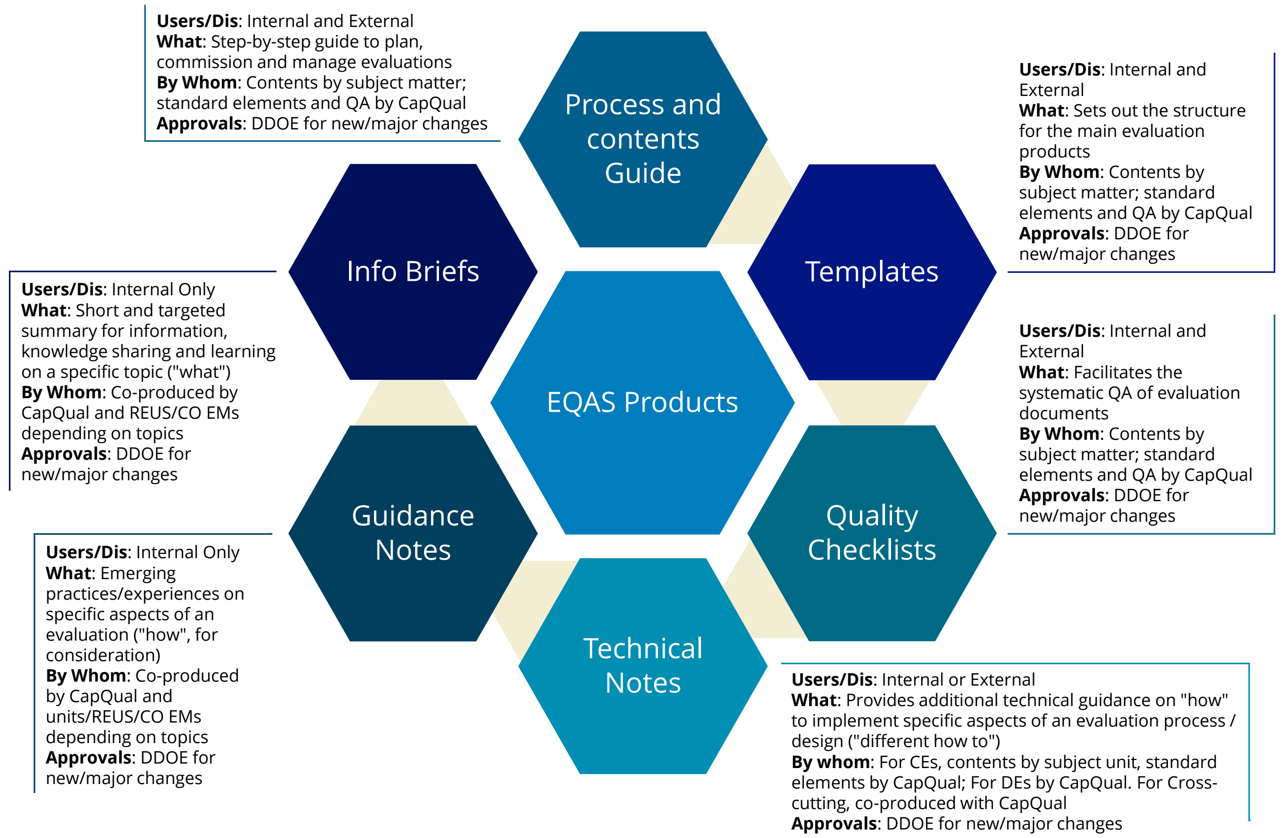
This Process Guide forms one element of the WFP Decentralized Evaluation Quality Assurance System (DEQAS). It should be used alongside the other elements summarized in the figure below, as well as:

- Decision tool for making the choice: decentralized evaluation or review? This is to be used to decide between a decentralized evaluation or a review based on the objective and expected use of the exercise.
- The mini guide for country directors/deputy country directors sets out the main roles and accountabilities of WFP management and the key decisions they have to take at each phase of a decentralized evaluation process. Typically, those are country directors, deputy country directors, regional directors, deputy regional directors, or directors of a headquarters division.

The Guidance Package is also part of a broader set of support mechanisms for decentralized evaluations, presented in Box 1 and further detailed in the WFP Support Mechanisms for Decentralized Evaluations brief, which include:

- an Evaluation Learning Programme (EvalPro): a comprehensive learning programme developed to strengthen the evaluation capacity of WFP staff. EvalPro targets all staff, providing them with the skills and knowledge they require to engage with or manage evaluations. If a staff member is about to embark on a decentralized evaluation and wishes to take the course on managing a decentralized evaluation the staff member would email Global.EvalPro@wfp.org to register;
- a decentralized evaluation outsourced quality support service (DEQS), which provides managers of decentralized evaluations with an impartial assessment of the quality of draft terms of reference (ToR), inception reports (IRs) and evaluation reports (ERs) along with practical recommendations to improve the draft evaluation deliverables. All draft ToR, inception reports and evaluation reports for decentralized evaluations must undergo a review by the DEQS. It is free of cost for the users of the service;
- a Contingency Evaluation Fund (CEF) to support country offices that adequately planned and budgeted for a decentralized evaluation but are facing genuine resource constraints;
- a help desk to complement the **support from regional evaluation units**, and provide technical advice and support to country offices, regional bureaux and headquarters divisions, on enquiries that might arise during the evaluation process;
- access to evaluation expertise through long-term agreements (LTAs) with evaluation service providers;
- a post hoc quality assessment (PHQA) system to provide an expert external appraisal of the extent to which all final evaluation reports meet evaluation quality standards; and
- dashboards to visualize data related to Office of Evaluation management information systems for evaluations, in terms of planning, monitoring and searching on budgets, duration, coverage, etc.

Figure 2: Overview of DEQAS guidance package



Box 1: WFP support mechanisms for decentralized evaluations

3. EVALUATION TYPES AND CATEGORIES IN WFP

WFP adheres to the United Nations definition of evaluation: "An assessment that is as systematic and impartial as possible. It analyses the level of achievement of both expected and unexpected results by examining the results chain, processes, contextual factors and causality using appropriate criteria such as relevance, effectiveness, efficiency, coherence, impact and sustainability of WFP's activities, strategies and policies, and their contribution to the development and humanitarian processes of countries that receive WFP assistance. An evaluation should provide credible, useful, evidence-based information that enables the timely incorporation of its findings, recommendations and lessons into the decision-making processes of organizations and stakeholders."[1]

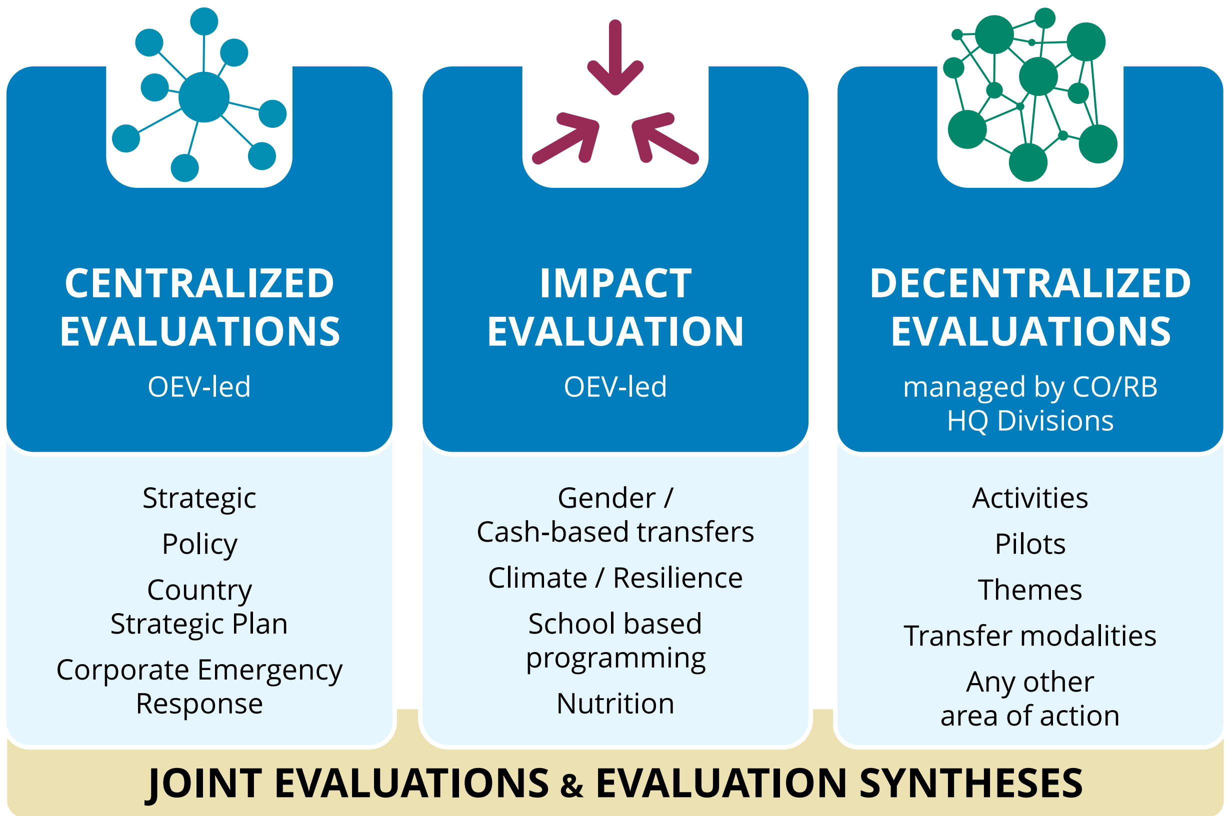
Evaluations have a different purpose and approach than other exercises such as needs assessments, appraisals, monitoring and audits. See the Technical Note Glossary of Terms for a list of definitions. Evaluations fall into three categories and under each category there are different types of evaluations (See Figure 3).

Centralized evaluations are commissioned and managed by the Office of Evaluation and presented to the Executive Board for consideration. They focus on corporate strategies and policies, global programmes, strategic issues and themes, corporate emergencies and CSPs. They take place at the global or the national level.

Decentralized evaluations are commissioned and managed by country offices, regional bureaux and headquarters-based divisions other than the Office of Evaluation and are designed to meet the needs of the commissioning units. They are not presented to the Board. They can cover activities, pilots, themes, transfer modalities or any other area of action at the subnational, national or multi-country level.

Impact evaluations are managed by the Office of Evaluation at the request of country offices. An impact evaluation is an assessment of the positive and negative, direct or indirect, intended or unintended changes in the lives of affected populations in receipt of WFP interventions. They measure changes in development outcomes of interest for a target population that can be attributed to a specific programme or policy through a credible counterfactual. Such evaluations demand a high standard of data, design and resources. They should only be considered where the need is clear, for example, the demand of evidence needed is high and the resources are available to undertake it successfully. They are usually undertaken during programme implementation over a multi-year period. Impact evaluations are not presented to the Executive Board. The Impact Evaluation Decision Guide helps determine whether a proposed evaluation fits with the definition of impact evaluation. The Office of Evaluation impact evaluation team and a regional evaluation unit (when an evaluation is commissioned in their region) provide advice on the feasibility and relevance of impact evaluations. For more information, see WFP's Impact Evaluation Strategy (2019–2026).

Figure 3: Evaluation types



All evaluations have the following commonalities:

- they contribute to both objectives of accountability and learning;
- they apply international evaluation criteria, as well as the United Nations System-Wide Action Plan (UN-SWAP) on gender equality and women's empowerment (GEWE), and the United Nations Disability Inclusion Strategy (UN-DIS) standards for evaluation;
- they must be conducted by independent consultants or firms; and
- findings must be publicly shared.

Types of decentralized evaluation

The nature of the subject of the evaluation will determine the type of decentralized evaluation, as follows:

- **Activity evaluation**: Assesses an ongoing or completed WFP activity, from design to implementation and results. It can cover one or several activities within a CSP but should not attempt to cover the entire portfolio. It can also cover one activity across several CSPs in multiple countries.
- **Pilot evaluation**: Assesses a pilot intervention, defined as an experiment or test, often small-scale at first, before introducing an intervention more widely. Evaluations of pilot projects generate evidence on: their relevance; results, whether intended or not; and how the pilot project has impacted target communities to determine whether the pilot can be scaled up in the same country or elsewhere and, if so, under which conditions. The evaluation of a pilot is critical prior to any scale-up or replication and before decisions are made on the design of a potential successor intervention.
- **Transfer modality evaluation**: Assesses, notably in view of beneficiary preferences, the appropriateness of the choice of a transfer modality(ies), defined as the mode in which assistance is transferred to intended beneficiaries (in-kind, commodity vouchers and cash). Transfer modality evaluations assess the intended or unintended results of various transfer modalities and their relative efficiency with the aim of understanding when, why and how a given transfer modality or combination of transfer modalities best achieves the desired outcomes compared to another.
- **Thematic evaluation**: Covers a theme across the entire CSP portfolio or selected activities in one or more countries, or globally. They provide a "big-picture" perspective on how WFP is performing; how the organization could further improve in a given thematic area, such as gender equality and the empowerment of women, capacity strengthening, and protection; and identify good practices in a given thematic area and within a range of operational contexts.
- **CSP outcome evaluation**: Assesses all ongoing or completed activities under a CSP outcome in one country. It should not attempt to cover the entire country office portfolio.

The Technical Note on Decentralized Evaluation Types clarifies specificities of the different types of decentralized evaluations, including: what they assess; how they are used; what questions they typically answer; and the specific data requirements.

Joint

As part of the 2030 Agenda for Sustainable Development (2030 Agenda), **joint evaluations** are increasingly encouraged as they offer opportunities to assess collective contributions to common goals in a more holistic manner, by learning together and strengthening partnerships. Joint evaluations certainly add value in cases of joint programmes: when commissioned jointly with government partners, they constitute an interesting opportunity to leverage national evaluation capacity strengthening. While offering numerous advantages, joint evaluations can be complex to manage and require a common understanding of quality standards, clear governance mechanisms and a common purpose. The Technical Note for Joint Evaluations includes a decision tree to facilitate decision making on whether to commission a joint evaluation.

Multi-country decentralized evaluations cover more than one country on a similar theme, using one evaluation team and one management and quality assurance process through evaluation process that produces a full evaluation report for each of the countries involved. There may or may not be a summary report that brings together common findings and lessons. This **information brief on multi-country decentralized evaluations** provides further information on its rationale, decision making and operationalization.

4. ROLES AND RESPONSIBILITIES

RB & HQ DE

Roles and responsibilities in a decentralized evaluation are described in the WFP Evaluation Charter and are further detailed below. For decentralized evaluations commissioned by a regional bureau or a headquarters division, the country director's main responsibilities apply to the regional/ deputy regional director or the headquarters division director. The regional evaluation unit responsibilities will apply to the Office of Evaluation's Capacity and Quality Unit (CapQual).

Joint

For multi-country decentralized evaluations, the decentralized evaluation would be co-managed by one evaluation manager per country and a member of the regional evaluation unit, steered by an evaluation committee made up of at least one representative from each country office and regional bureau.

In the context of a joint evaluation commissioned with partners, "country office management" should be interpreted as "management of the various commissioning entities".

Commissioning office management

The role of country director/deputy country director, headquarters division director (for headquarters-commissioned decentralized evaluations) or regional bureau regional director or deputy regional director (for regional bureau-commissioned decentralized evaluations) is to:

- decide strategically on the number of decentralized evaluations, their respective scope, type, budget and timing at CSP formulation stage and ensure adequate sequencing with other evidence generation exercises;
- budget adequately in the country portfolio budget and the evidence planning and budgeting tool (EPB). The average to date is USD 130,000 for the conduct of a decentralized evaluation;
- ensure that the evaluation is independent and impartial as per the Evaluation Policy's provisions and is free from undue influence, and that reporting is unbiased and transparent. This implies nominating the evaluation manager in line with guidelines (see Step 1.5 below); setting up an evaluation commission (EC) and an evaluation reference group (ERG); ensuring that the evaluation is conducted by qualified independent consultants who sign the Pledge of Ethical Conduct and Confidentiality Agreement; ensuring that the required information is provided to the evaluation team; and discussing with country office staff the implications of impartiality and independence principles;
- ensure that evaluation plans are reflected in the country office workplan and individual performance assessment and competency enhancement (PACE); and that the evaluation manager has a realistic workload that allows him/her to devote sufficient time to the management of the evaluation as well as complete EvalPro 4;
- undertake the online course EvalPro 2 evaluation for decision makers;
- provide strategic steering of the evaluation process and make key decisions on the evaluation, notably: budget, funds allocation and selection of the evaluation team; approve the ToR, and inception and evaluation reports;
- lead preparation of a management response and clear it (See Technical Note on Management Response to Decentralized Evaluations Recommendations);
- ensure that the final evaluation report and management response are published and widely disseminated;
- nominate a staff member responsible for tracking implementation of follow-up actions; and
- use evaluative evidence to inform the design and/or implementation of the CSP or other initiatives.

The evaluation manager (EM)

Members of the monitoring and evaluation (M&E) team are the primary candidates to take on the role of evaluation manager (see Step 1.5 on evaluation management options) but the director of the commissioning office may decide to nominate another staff member. The evaluation manager is responsible for the evaluation process through all phases and:

- enrols and learns through the Evaluation Learning Programme (EvalPro 4 managing a decentralized evaluation);
- prepares the ToR;
- leads assessment of proposed evaluators in collaboration with procurement or human resources (HR) units and the regional evaluation unit;
- presents the budget and proposed evaluation team to the evaluation committee for review, and to the procurement authority for approval;
- is the main contact point in WFP for the evaluation team leader throughout the evaluation process; maintains regular communication; and provides updates on the evaluation to key stakeholders;
- convenes, for the evaluation chair, the evaluation reference group and the evaluation commission;
- consolidates the evaluation document library;
- coordinates the administrative and logistical support for the evaluation;
- comments on and quality assures the evaluation products in compliance with DEQAS;
- submits final, quality assured ToRs, inception reports and evaluation reports for approval of the director of the commissioning office;
- supports the commissioning office management in the development of a management response; and
- disseminates the evaluation report and management response widely and requests that the Office of Evaluation publish them on WFP websites.

Monitoring and evaluation team

Along with the programme team, M&E staff help identify knowledge gaps, support appropriate planning and budgeting for evaluations and provide advice to country office management accordingly. Other roles and responsibilities of M&E staff are outlined in the functional summary and include:

- ensuring that decentralized evaluation plans are embedded in the country portfolio budget and reflected in the country office evidence planning and budgeting tool (EPB) (at CSP formulation stage);
- contributing to the evaluability[2] of WFP undertakings – establishing baseline information, performance indicators and targets for expected results as well as providing latest logframes/monitoring data;
- clarifying the type of information/evidence that monitoring systems and processes can provide and contribute to discussions on when the evaluation results would be most useful for decision making and action;
- providing the evaluation manager with relevant monitoring reports and data on performance indicators for measuring the achievement of the subject being evaluated;
- providing support in reviewing and quality assuring evaluation products;
- engaging with the evaluation team as relevant and supporting the overall evaluation process;
- ensuring that baselines and monitoring data are available to enable evaluation;
- seeking clarity on requirements for ensuring impartial and credible evaluation; and
- contributing to the preparation of the management response and implementation of follow-up actions.

Procurement team

The role of procurement staff in evaluation includes:

- advising the evaluation committee on procurement options (long-term agreement or competitive tender), referring to appropriate procurement manuals and the Technical Note on Options for Contracting Evaluation Teams;
- leading communication with long-term agreement firms, participating in the assessment of proposals, and managing the mini bid with long-term agreement firms (in offices outside headquarters);
- advising the evaluation committee on the implications of the contracting modality on time and the level of effort required to enable an estimation of the evaluation timeline;
- participating in the assessment of evaluation proposals submitted by the evaluation firms;
- reviewing the draft ToR, inception and evaluation reports and providing comments, focusing on the analysis and presentation of procurement-related data to ensure they are factual; and
- contributing to the preparation of the management response and implementation of the follow-up actions.

Human resources (HR) team

When the decision is made to contract individual evaluators, HR staff have an important role to play to ensure that contracting is done according to appropriate procedures. Details on the roles of HR staff during the different phases of a decentralized evaluation can be found in the Evaluation Function Summary for HR Staff. These include:

- advising the evaluation committee on procurement options (long-term agreement or competitive tender), referring to appropriate procurement manuals and the Technical Note on Options for Contracting Evaluation Teams;
- leading communication with long-term agreement firms, participating in the assessment of proposals, and managing the mini bid with long-term agreement firms (in offices outside headquarters);
- advising the evaluation committee on the implications of the contracting modality on time and the level of effort required to enable an estimation of the evaluation timeline;
- participating in the assessment of evaluation proposals submitted by the evaluation firms;
- reviewing the draft ToR, inception and evaluation reports and providing comments, focusing on the analysis and presentation of procurement-related data to ensure they are factual; and
- contributing to the preparation of the management response and implementation of the follow-up actions.

Budget programme/finance team

The role of budget programme and finance staff in evaluation include:

- advising country office management on the budget available;
- contributing to the planning and budgeting processes;
- advising the evaluation committee on the availability of funds under the budget lines intended to fund the evaluation, in accordance with the Technical Note on country-specific evaluation planning and budgeting;
- providing finance-related documents to be used in preparation of the ToR;
- providing finance data and reports, and any additional data/information or clarification required;
- reviewing the draft evaluation report and providing comments, focusing on the analysis and presentation of finance and resource management data to ensure they are factual; and
- contributing to the preparation of the management response and implementation of follow-up actions as required.

Government partnership officers

The role of government partnership officers in relation to evaluation are outlined in the Evaluation Function Summary for Government Partnership Officers. These include:

- engaging with donors as relevant to manage expectations on evaluations, ensuring a common understanding of the process, with a view to ensuring WFP evaluation coverage norms are applied;
- should a donor request explicit provisions for evaluation in a grant/contribution agreement, seeking to secure a sufficient evaluation budget;
- providing the evaluation team with relevant background information on partnerships and donor engagement;
- briefing the evaluation team on specific issues related to partnerships;
- reviewing the draft ToR, inception and evaluation reports and providing comments, focusing on the analysis and presentation of partnership data; and
- contributing to the preparation of the management response and implementation of follow-up actions as required.

Vulnerability analysis and mapping (VAM) team

The role of VAM staff in evaluation are outlined in the functional summary. These include:

- providing the evaluation team with relevant background information on food security and needs assessments reports for the activities/ interventions;
- clarifying the type of information that vulnerability analysis and mapping can provide;
- contributing to briefing the evaluation team during the inception and data collection phases on food security assessment systems, processes and reports;
- reviewing the draft ToR, inception and evaluation reports and providing comments, focusing on the analysis and presentation of food security data; and
- contributing to the preparation of the management response and implementation of follow-up actions as required.

At the **regional level**, roles and responsibilities in a decentralized evaluation commissioned by a country office are described in the Evaluation Charter, the Evaluation Function Terms of Reference and are further detailed below:

The regional director (RD)

The regional director is overall accountable for steering the evaluation function across the region and:

- applies the impartiality and coverage provisions for decentralized evaluations as per the Evaluation Policy;
- engages in regular consultations with the Office of Evaluation and country offices to ensure complementarity among evaluations commissioned across the function;
- ensures that plans for evaluations are included in the design of regional strategies, interventions and other initiatives;
- embeds roles and accountabilities for evaluation in country directors' performance management systems;
- ensures that resources are budgeted for and allocated to the management of independent evaluations, including those commissioned by country offices, and provides regional-level support and oversight;
- oversees application of evaluation quality assurance procedures;
- ensures that management responses to evaluations commissioned by regional and country offices are prepared and made publicly available and that relevant follow-up actions are undertaken; and
- ensures that CSPs and new programmes, initiatives and strategies prepared in the region are based on evidence from evaluations.

Regional evaluation unit

The core purpose of the regional evaluation units is to support country offices and regional bureaux in commissioning, managing and using decentralized evaluations, as well as in using evidence from other types of evaluations. The regional evaluation unit:

- supports the regional director in his/her evaluation responsibilities, listed above;
- supports regional bureau and country office management to operationalize the impartiality provisions;
- supports appropriate planning and budgeting for evaluations in consultation with budget and programme staff;
- ensures that staff managing and/or steering decentralized evaluations are enrolled in the Evaluation Learning Programme;
- acts as first port of call for advice/support to country offices during planning and conduct of decentralized evaluations;
- reviews the management response to decentralized evaluations in collaboration with the regional monitoring adviser, the senior regional programme adviser and other regional bureau colleagues and technical staff concerned and assures the regional bureau management that all recommendations have been sufficiently responded to before the regional director's endorsement of the management response;
- ensures that all decentralized evaluation reports and management responses are publicly available; and
- promotes the use of evaluation findings in CSP design and implementation.

Other regional advisers (monitoring, school feeding, social protection, country capacity strengthening etc)

Depending on the subjects of evaluations, different regional advisors play certain roles including:

- participating in evaluation reference groups for the decentralized evaluation and commenting on draft evaluation products;
- providing inputs to the management response when recommendations are directed to their areas of work; and
- as part of the the regional monitoring unit, being responsible for uploading management responses for decentralized evaluations in the R2 system and periodically liaising with country offices and regional bureau units on implementation of follow-up actions.

At the **headquarters level**, roles and responsibilities are described in the Evaluation Charter, the Evaluation Function Terms of Reference and are further detailed below:

Office of Evaluation (OEV)

The Office of Evaluation is responsible for setting overall standards and the normative framework for evaluation across the organization including:

- defining the norms and standards, safeguards for impartiality, and expected coverage for decentralized evaluations;
- periodically updating the DEQAS based on feedback by users and changes in the normative framework;
- providing technical support and guidance to headquarters divisions, regional bureaux and country offices commissioning decentralized evaluations through the Office of Evaluation help desk;
- supporting appropriate planning and budgeting for evaluations in line with appropriate budgeting and programming guidelines;
- managing a range of support mechanisms, including an outsourced quality support service; the Contingency Evaluation Fund; and long-term agreements with firms providing evaluation services;
- acting as first port of call for advice/support to regional evaluation units and headquarters divisions during planning and the conduct of decentralized evaluations;
- ensuring independent post hoc quality assessment of all final evaluation reports and sharing the results with the commissioning offices;
- on behalf of commissioners, publishing the evaluation products and management response on WFP evaluation internal and external websites; and
- promoting the use of evaluation findings in CSP design and implementation.

Other headquarters divisions

Headquarters divisions may commission and manage decentralized evaluations, with support of the Office of Evaluation. For decentralized evaluations commissioned by regional bureaux or country offices, depending on subjects of evaluations, different headquarters divisions play certain roles including:

- providing technical support and advice on key policy, strategic and programmatic considerations, relevant to the subject of evaluation, to ensure they are understood from the onset of the decentralized evaluation, as required;

ToC

- clarifying requirements for donor-demanded decentralized evaluations and acting as liaison between regional bureau/country office and donors as appropriate;
- participating in some evaluation reference groups for decentralized evaluations as relevant;
- providing inputs to the management response when recommendations are directed to their units; and
- using decentralized evaluation findings for the development of policies, strategies and guidance and sharing lessons

External stakeholders

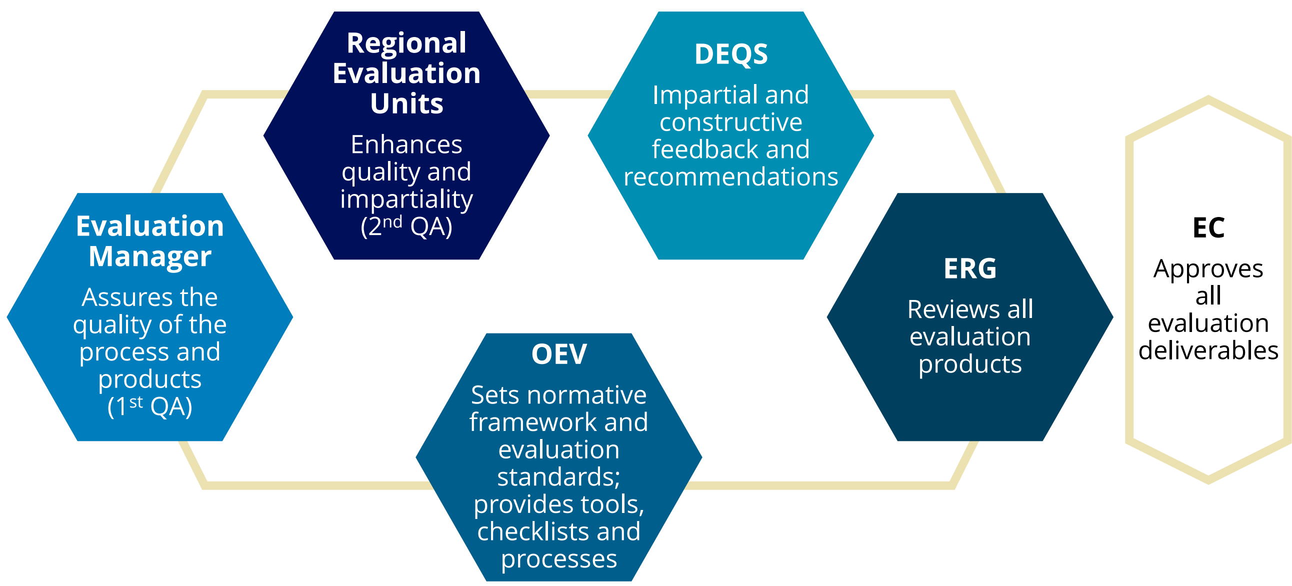
Many external stakeholders such as beneficiaries, governments, other humanitarian and development actors, such as national and international non-governmental organizations (NGOs), donors and United Nations agencies will have a direct or indirect interest in an evaluation. Assessing their level of interest (stakeholder analysis) and planning for their involvement in the evaluation is one of the early steps to be undertaken. Their degree of involvement will vary, including:

- key informants;
- participating in an evaluation reference group;
- reviewing and commenting on draft deliverables; and
- providing insights/information and follow-up on recommendations as appropriate.

The evaluation team

The team is composed of external evaluators who have no prior association with the subject of the evaluation. The team is expected to conduct the evaluation in adherence to DEQAS standards, the [UNEG Norms and Standards] and the [UNEG 2020 Pledge of Ethical Conduct in Evaluation]. The team is responsible for designing the evaluation methodology in line with the ToR, undertaking the data collection, analysis, drafting the report, presenting findings and recommendations to the commissioning office to get feedback, and finalizing the report based on feedback. The team is expected to consider gender, disability and broader inclusion issues throughout the evaluation process and adequately address them in the report.

Figure 4:. Quality assurance process for country office-led decentralized evaluation: roles and responsibilities



5. QUALITY ASSURANCE

Quality assurance cuts across all phases of a decentralized evaluation as it includes assuring process adherence to norms and standards and clearance of all evaluation outputs. For the preparation phase, the WFP evaluation manager has the primary responsibility for developing a ToR in line with the DEQAS. For inception and reporting phases, the evaluation team leader and, when relevant, the evaluation firm that hired her or him, have the primary responsibility for timely delivery of evaluation reports that meet WFP evaluation quality standards. The evaluation manager is responsible for ensuring that the evaluation progresses, as per the DEQAS Process Guide, and for conducting a rigorous quality control of the evaluation products ahead of their approval by the director of the commissioning office.

- **First level of quality assurance (QA1)** is undertaken by the evaluation manager for all agreed evaluation deliverables produced by the evaluation team at each stage of the process: each draft of inception and the evaluation reports.
- **Second level quality assurance (QA2)** is undertaken by the regional evaluation unit on all agreed evaluation deliverables produced by the evaluation team. In the case of regional bureau-led and headquarters decentralized evaluations, QA2 is done by the Office of Evaluation's capacity and quality unit.
- **Final approval of documents**. All documents to be approved by the director/deputy director of the commissioning office, who chairs the evaluation committee.

RB & HQ DE

When conducting quality assurance, it is important to pay attention to the tone and clarity of comments to avoid any misunderstanding or contradictions. If relevant, make a distinction between comments calling for imperative action, and those expressing reservations and/or requiring clarification. Table 2 below provides some examples of what type of inputs are considered appropriate or not during the quality assurance process. When the feedback provided is not appropriate or clear, the evaluation team should take note of it in the evaluation comments matrix and, as needed, request additional clarifications or evidence.

Table 2: Examples of quality assurance inputs

Comment type	Appropriate	Not appropriate
Factual correction with supporting evidence	√	
Supplementary information (within timeframe) that changes analysis	√	
Rephrasing that adds clarity/conciseness/ appropriateness for the organization	√ (if logical and does not change meaning)	If it changes meaning or intended message
Higher level (more in-depth) analysis than is in current draft	√ (if supported by clear reasoning and logic)	
Request for additional supporting evidence to back findings and conclusions stated	√ (if reasonable)	
Editorial comments	√ Only when helping to keep within word limit, ensure use of appropriate terminology and manage readability	Grammatical comments are not appropriate as the report should be edited later
Supplementary information (outside timeframe or scope) that changes analysis		√ Leads to scope creep
Diverging interpretation of evidence		√ (unless supported by additional in-timeframe evidence)
Commentator opinion		√
Proposed amendments without clear reasoning/case for treatment		√
Subjective editing style		√

ToC

Introduction

The planning phase has the following purposes:

- To ensure that decentralized evaluations are planned with a clear intended use and are embedded in the WFP programme cycle, thereby contributing to evidence-based programme design and implementation, fundraising and advocacy.
- To ensure adequate and balanced evaluation coverage, which involves determining in advance the type and number of evaluations needed as well as broad parameters for their respective purpose, scope and timing.
- To ensure that evaluation costs are adequately budgeted for in country portfolio budgets (CPBs) under appropriate cost categories covering staff, contracted services and other costs:
  - CSP evaluations and thematic decentralized evaluations budgeted under direct support costs (DSC);
  - activity-specific decentralized evaluations budgeted under implementation costs; and
  - national evaluation capacities development activities budgeted under country capacity strengthening.
- To ensure that all evaluations are reflected in the country office's assessment, monitoring and evaluation (AME)/evidence planning and budgeting (EPB) tool,[3] building synergies and the right sequencing across the various assessments, monitoring and evaluation exercises including baselines, outcome monitoring, reviews and studies.
- To ensure that evaluation is on the agenda and office and individual work plans and that the necessary human resources and time are allocated.

The WFP minimum evaluation coverage norms for decentralized evaluations:

- **Country office**. At least one decentralized evaluation (For example, activity or thematic evaluation or CSP strategic outcome evaluation) per country office per interim I/CSP cycle. The number, timing and scope of decentralized evaluations are defined by country offices based on WFP learning needs and donor and/or partners' learning, and accountability needs. Larger country offices may commission more than one decentralized evaluation.
- **Regional bureau**. No specific norms but criteria to guide decision making on evaluation (see Box 2) should be applied, particularly for multi-country evaluations. Other criteria may include projects implemented through funds channelled through or managed by the regional bureau.
- **Headquarters office/division**. No specific norms but criteria to guide decision making on evaluation should be applied (see Box 4). Other criteria may include projects implemented through funds channelled through or managed by headquarters or areas of corporate learning that are not covered by strategic centralized evaluations.

There are multiple sources of demand for decentralized evaluations, such as:

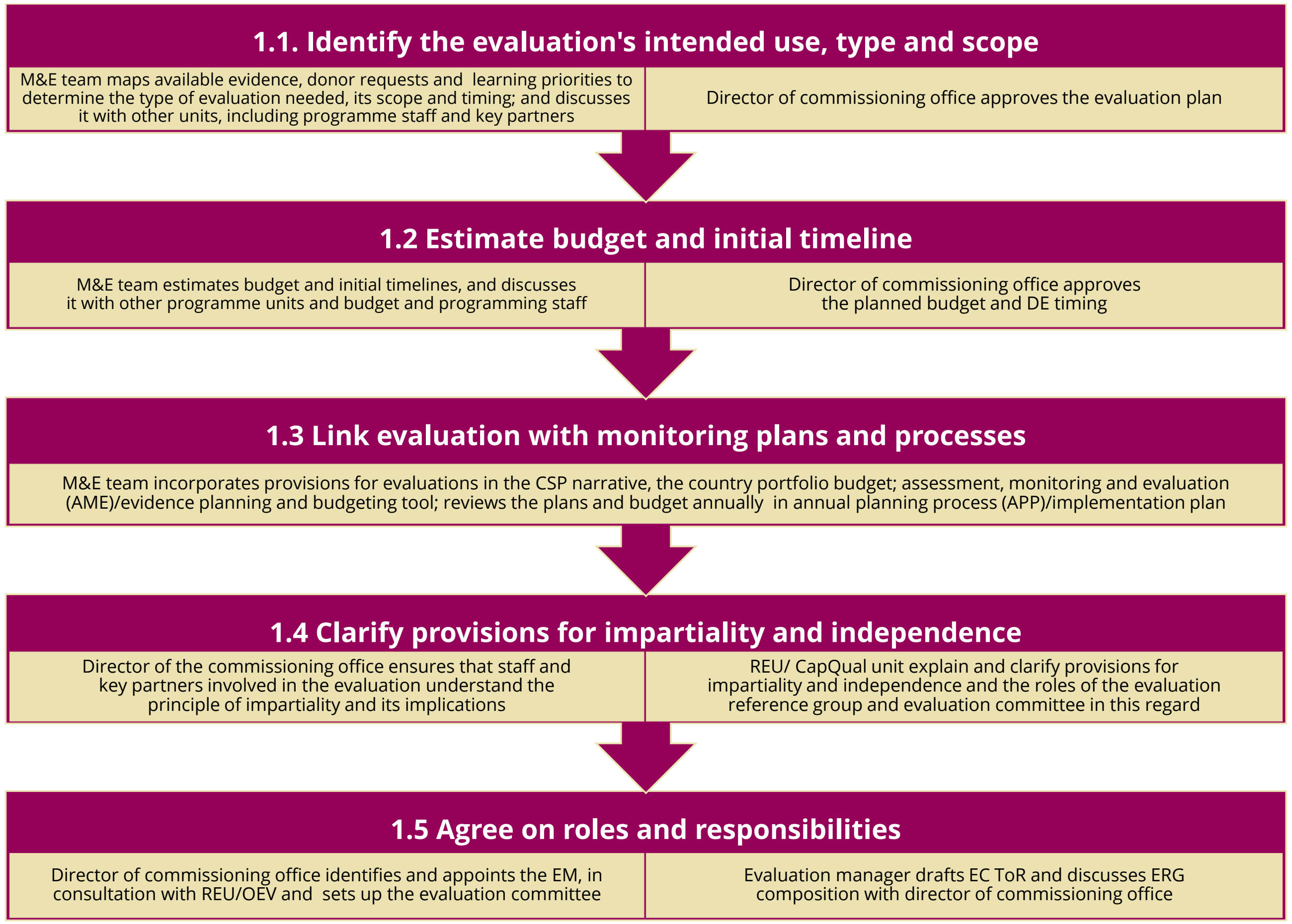
- country office/regional bureau/headquarters units management's need for enhanced knowledge to improve future programme design and implementation, advocacy strategy, fundraising, etc.; and
- requests from donors and partners to generate evidence and demonstrate results, including through joint evaluations.

When deciding whether to conduct a decentralized evaluation, the criteria presented in Box 2 could be used to guide the decision making process.

Box 2: Criteria to guide decision making for decentralized evaluations

- Strategic relevance to WFP
- Evidence gaps (at the country, regional or global levels)
- Level of programme expenditure
- Scale of emergency response
- Before replication or scale-up of pilots, innovations and prototypes
- Innovative results (for example, achieved across a region or through innovative multi-country programmes that are centrally funded or supported)
- Formal commitments to stakeholders (for example, to national partners to inform national programmes, or to funders as part of funding requirements)
- Likelihood of influencing policymaking or potential for leveraging partnerships
- Feasibility of undertaking the evaluation

Figure 5: Summary of key responsibilities – planning phase[4]



ToC

1.1. Identify the evaluation's intended use, type and scope

What could go wrong?

Lack of willingness by senior management due to timing, budget, complex processes, limited human capacity or perception of 'too many evaluations' or being 'over evaluated'

Too many evaluations (due to donor demands) or none at all are planned.

Good practices

Ensure evaluations are included in the CSP at the design stage and adequately sequenced with other evidence-generation exercises, included the CSP evaluation and the mid-term review

Identify the right evaluation types and tools through a thorough review and learning needs based on past evaluation evidence and nature of programming. Avoid 'compliance/norms' orientation

Advocate for more involvement of senior management in evaluation processes, using the regional evaluation committee (REC) as a platform for advocacy and escalating when needed.

Planning for an evaluation should be done when formulating a CSP at the country office level. Initial decisions (even if broad and tentative) contribute to a more coherent and timely evaluation plan as well as early resourcing. The Technical Note on country-specific evaluation planning and budgeting (covering centralized, decentralized and impact evaluations) guides country office and regional bureau staff involved in the design of CSPs as well as the preparation of the country portfolio budget.

The multi-country decentralized evaluations (MCDEs) approach should be considered and discussed by the regional evaluation committee, informed by background work done by the regional evaluation unit on when and where it may be feasible to cluster decentralized evaluations to fill evidence gaps, meet coverage norms and meet donor requirements while contributing to the evidence base around identified themes. The regional evaluation unit should discuss and advise whether proposed MCDEs should be included in the regional evaluation plan (REP) of a given year.

Intended use

The head of the M&E unit should assemble the necessary information including a mapping of available evidence (areas covered by recent evaluations); strategic learning priorities identified by regional bureaux and country offices; evaluation commitments made with donors/ partners; and recommendations from past evaluations and reviews, etc.

In a second step, s/he should facilitate open discussions with programme and operational teams,[5] country office management and key partners around the following basic but fundamental questions on the intended use of the evaluation:

- What evaluation evidence is currently available? Where are there gaps? what are the learning priorities?
- What is the main purpose of the evaluation? How will the evaluation be used? Which decisions could be informed by this evaluation? When are those decisions being made?

Box 3: Meeting donors' accountability and learning needs

- Different arrangements can be made for evaluations to meet donors' accountability and learning needs regarding their contributions to WFP. In the framework of its coverage norms, WFP does not make special arrangements for evaluating every individual donor's specific contribution to a WFP intervention, but interventions may be subject to evaluation as part of broader, strategic evaluations.
- Should donors' own accountability requirements with regard to taxpayers not be compatible with the principle of donor harmonization, donors may request that WFP conducts an evaluation of their specific contributions. Decentralized evaluations are typically used to meet these increased accountability requirements of some donors, even if the country office has already met its coverage norm by commissioning another decentralized evaluation. In these cases, the commissioning unit should manage donor expectations on what they can expect from a decentralized evaluation with allocated budget and clarify with the donor whether a decentralized evaluation or a review would meet donor needs. If an evaluation of a donor grant is foreseen, it should be explicitly mentioned in the funding agreement and adequately budgeted for, using the average of USD 130,000 as a guide.
- Should all the above options not be possible, donors may decide to commission their own external evaluation. Such evaluations often focus on the donor's broader support to a given country/region or initiative, for which WFP is one of many stakeholders. Any donor-led evaluation should be formalized in the relevant partnership agreement between the donor and WFP. WFP should be informed at the outset when a donor intends to commission its own evaluation and be given the opportunity to engage meaningfully at different stages of the evaluation process.

- Which key questions should be addressed?
- Whose needs and interests will it serve? Are there other interested actors in this sector – government, donors, or local or national groups – who would inform or could be informed by this evaluation?
- Would there be value in jointly commissioning the evaluation with partners? Who would these be?

Addressing these questions helps to determine the most suitable type of evaluation needed, its scope and timing, even if these may be tentative and will evolve over time.

Type and scope of evaluation

The broad parameters of the evaluation scope (what will be included/excluded) should be determined by the programme/operational and M&E teams during the planning phase. This includes: (i) the activity(ies)/theme/ transfer modalities to be evaluated, which will help to identify the evaluation type as defined in Section 1.3. and the Technical Note on Decentralized Evaluation Types; (ii) the period to be covered; and (iii) the geographical area to be covered.

Decisions on the scope of a decentralized evaluation should be driven by the learning agenda, with the only exception being where WFP has committed to a specific donor to commission an evaluation of a given activity.

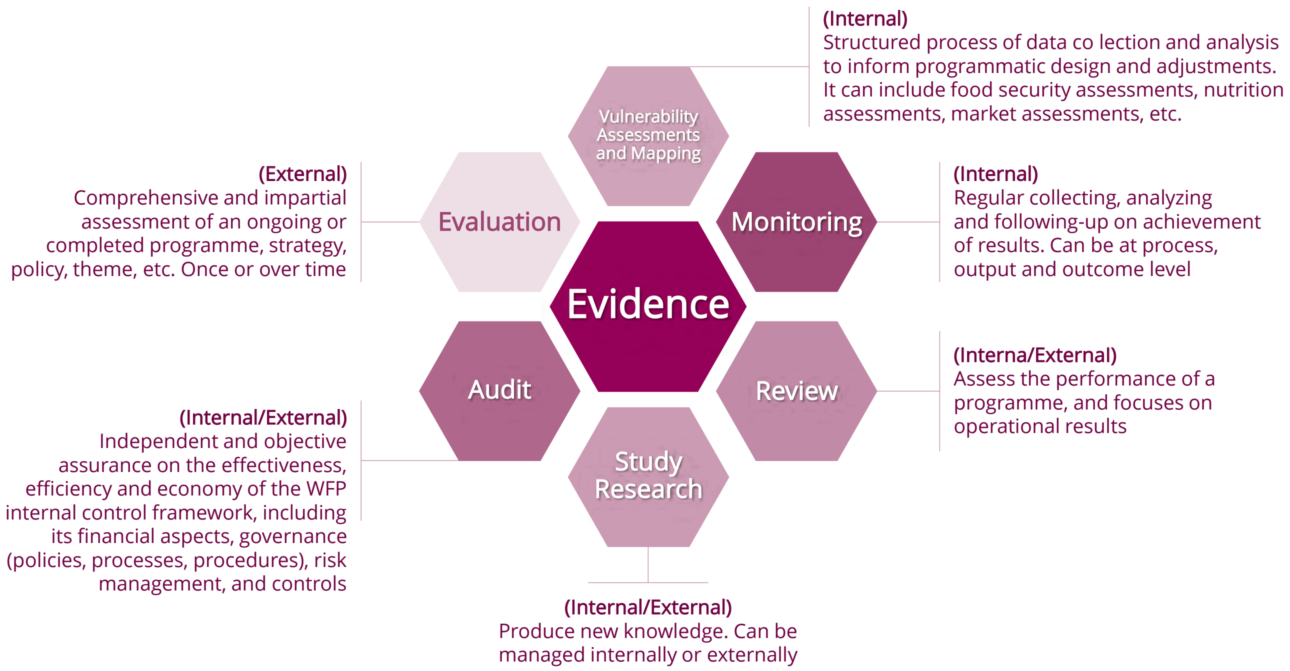
Evidence generation

As highlighted in WFP's Strategic Plan (2022-25), As highlighted in the WFP Strategic Plan (2022-2025), evidence is one of the seven corporate enablers, guiding and informing the WFP programmatic approaches and helping to demonstrate long-term impact. Robust, timely and relevant evidence is critical throughout the programme lifecycle of country strategic plans, from upstream assessments for programme design and targeting, to downstream monitoring and evaluation of efficiency, effectiveness and impact. Evidence is used for strategic and operational decision making, resource mobilization and accountability to stakeholders including beneficiaries and donors.

Different types of evidence within WFP can be generated through diverse types of tools including, evaluation, review, monitoring, assessment, studies/research or audit (see Figure 6). It is important to clarify the differences between evidence generated through these tools and the unique nature of evaluative evidence, which is generated independent of those involved in the design and implementation of operations.

The scope of a decentralized evaluation should be defined avoiding overlaps and ensuring optimal complementarity with other exercises such as the CSP mid-term review, the CSP evaluation and other reviews. Regional evaluation units can support the sequencing of these exercises in a way that reduces the burden on the country office team and maximizes the utility of the exercises. By optimizing synchronization in timing and scope, CSP evaluations will increasingly build on the evidence generated through decentralized evaluations and mid-term reviews as illustrated in Table 3 below.

Figure 6: Evidence generation within WFP



There is often no 'right' or 'wrong' choice between a decentralized evaluation or programmatic review although the orientation note on making the choice between a decentralized evaluation or a review can help to make a decision. In many situations – though not all – both are possible. The selection should be shaped by seven triggers: (i) demand; (ii) need; (iii) use; (iv) coverage; (v) timing; (vi) context; and (vii) duration and cost.

Decisions on whether to commission a decentralized evaluation or a programmatic review should:

- take into account the different features in terms of accountability and enhancing WFP credibility (a decentralized evaluation's added value) versus speed/duration and internal reflection (a programmatic review's added value);
- reflect the best option for the circumstance, given the varied operating conditions of WFP programming; and
- have the intention of maximizing value for the resources invested, based on the needs and expectations.

Table 3. Complementarities between DE, reviews, Mid-Term Reviews and CSPEs

	Decentralized evaluation	Programmatic review	CSP mid-term review	CSP evaluation
Purpose	Address specific areas of weak evidence and/ or donors' accountability requirements	Learning-oriented, geared towards operational decision making	Assess CSP performance of implemented activities to enable necessary adjustments	Assess the strategic positioning, coherence and performance of the WFP CSP
Scope	Specific activities/themes/components of the CSP		Entire CSP	
Timing	When needed		At mid-term	Penultimate year of CSP

1.2. Estimate initial budget and timeline

What could go wrong?

The decentralized evaluation is under-budgeted.

The decentralized evaluation timeline is not realistic, for example, too short to allow proper consultation with stakeholders and collection and analysis of data in ways that are credible.

Unrealistic expectations on what a decentralized evaluation can deliver (show/say) especially about IMPACT.

Good practices

Consult average rates of decentralized evaluation budgets in the region and for similar evaluation scope.

Review average timelines for decentralized evaluation and discuss the key factors affecting the duration including responsiveness.

Manage expectations, especially on what it takes to for a decentralized evaluation to assess impact (use examples of past decentralized evaluations).

Evaluation budget

The focus of the M&E team at CSP formulation stage should be on budgeting adequately for all planned evaluations for inclusion in the country portfolio budget and relevant evidence planning and budgeting (EPB) tools.[6]

The WFP Evaluation Policy does not specify the share of budget to be allocated to evaluation, as the total cost of a decentralized evaluation will depend on the type, scope, methodology and country context. However, experience to date shows that WFP decentralized evaluations cost on **average USD 130,000, noting that budgets typically range from USD 90,000 to USD 250,000**. At the minimum, key decentralized evaluation costs to be budgeted for include:

- **evaluation management costs (staff)**: minimum one third of the cost of a national officer (NOC) for one year – under direct support cost (DSC) staffing costs. Evaluation management costs relate to the staff time of the WFP staff member designated as evaluation manager;
- **evaluation conduct costs (contracted services)**: on average USD 130,000 – to evolve based on the considerations listed above – under the implementation costs of the activity(ies) being evaluated (budget line: "Evaluation Contracted Services") for activity "decentralized evaluations". If the decentralized evaluation cannot be attributed to specific activities (for example, thematic decentralized evaluation on gender), its costs should be budgeted under direct support cost;
- **others costs** to cover any elements for which WFP will pay for directly (not through contractors) such as hiring meeting venues, dissemination workshops etc.[7]

Figure 7. Key elements to consider estimating an evaluation budget

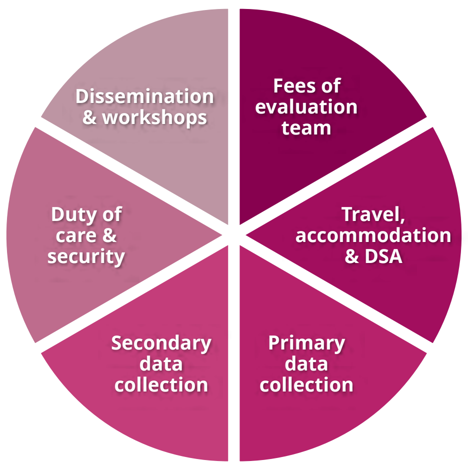
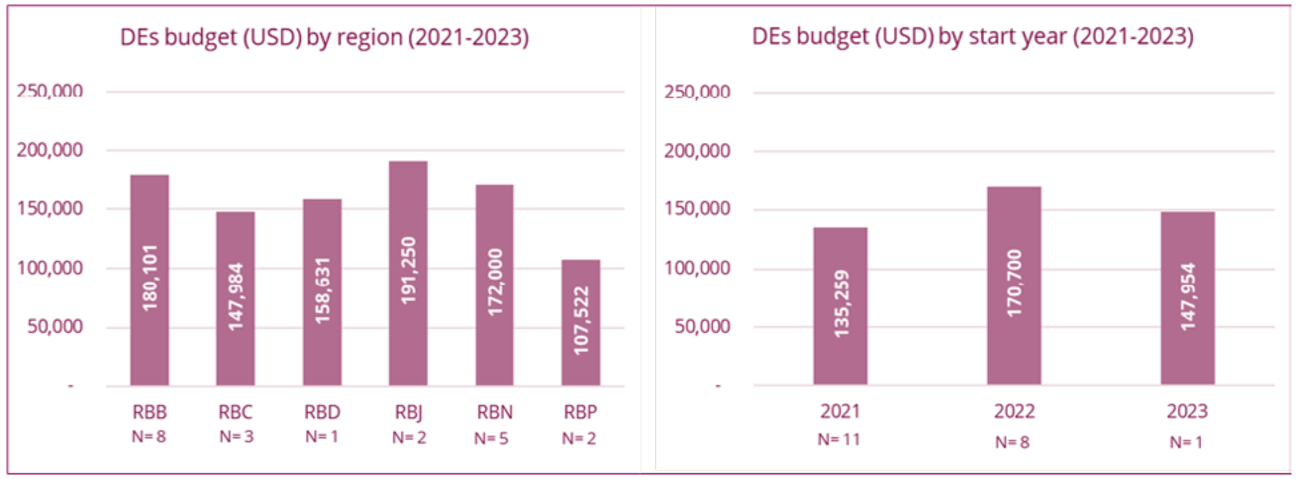


Figure 8. Key elements to consider estimating an evaluation budget



Note: excluding joint DEs.

When costing for a decentralized evaluation, some key cost drivers should be considered and adjusted based on budget situations:

- What type of evaluation is envisaged?
- What scope is envisaged in terms of number of activities, geographical coverage, and depth of analysis?
- What methodology do you foresee?
- What evaluation team size do you foresee? Are they national or international or a mix?
- What type of contracting process is envisaged (WFP long-term agreement, tender, direct consultancies)?
- Is additional collection of quantitative data required to answer the evaluation questions? If so, how many rounds of data collection are envisaged? What is the sample size?
- What are the costs for recruiting enumerators or traveling in the country?
- Do you foresee high security costs?
- Is additional translation required?
- How many dissemination workshops do you foresee with internal and external stakeholders? Are they face-to-face or virtual or hybrid? Who will meet the costs of attendance for different stakeholders?

Budget analysis of decentralized evaluations between 2021and 2023 shows some differences by region and start year as illustrated in Figure 8 below. Comparing between MGD/USDA-funded school feeding evaluations and other decentralized evaluations, the average budget for MGD/USDA decentralized evaluations was USD 212,248 and the non-MGD/USDA decentralized evaluations was USD 133,330. As for the programme area, the highest budget was for emergency preparedness.[8]

RB & HQ DE

Regional bureau-led decentralized evaluations conducted so far (n = 4) are more costly than the average cost of country office-led decentralized evaluations, averaging USD 146,000. The regional bureau decentralized evaluations have covered multiple countries as case studies but are not multi-country decentralized evaluations (MCDEs) as per the definition in this **Information brief on MCDEs**. While for MCDEs contracting is done at the regional level just like regional bureau decentralized evaluations, MCDEs use country office budgets and are co-managed by the regional bureau and country offices. Each country office contribution would be smaller than the average decentralized evaluation costs. The Technical Note on country-specific evaluation planning and budgeting provides further details on the key elements to consider when estimating the overall decentralized evaluation budget.

Evaluation timeline

In consultation with the programme team, the M&E team should determine an indicative timeline for the evaluation based on the following considerations:

- Has the activity been operating for a long enough period to create sufficient evidence for an evaluation?
- When will the key decision making processes that the evaluation should inform take place (for example, the design of the next phase of the intervention, decision on funding continuation etc)?

- Is it feasible to synchronize with monitoring plans/schedules to yield data for the evaluation without affecting the timeliness of evidence? For example, if outcome surveys are planned, it would make sense for these to be completed before the decentralized evaluation (see also Step 1.3 on linking with monitoring plans and processes).
- Is the time of the year significant, for example, during the school term or during a particular season, when physical accessibility and availability of stakeholders are critical.

The expected duration to deliver a decentralized evaluation (from preparation until evaluation report is approved) is estimated 7.5 months (30 weeks). In practice, decentralized evaluations are lasting up to 13.5 months. Table 4 presents the estimated/ideal as well as the actual duration of WFP decentralized evaluations from the preparation to the dissemination phase, based on an analysis of WFP 2021-2023 decentralized evaluation data. With this, and working backwards from the required completion date, the M&E team can determine an approximate start date.

Different factors contribute to delays at each phase. At the preparation stage, the evaluation manager may face administrative challenges with contracting the evaluation team or delays in decisions on the scope of the evaluation. At the analysis and reporting stage, inputs to the draft report may be delayed if key stakeholders are not available or responsive.

From past experiences, how long decentralized evaluations take does not seem to be significantly affected by the contract modality even although decentralized evaluations using open tenders and individual consultants take a slightly shorter time than those using long-term agreements (see Figure 9). Recruiting evaluation teams through long-term agreements should not take longer given that the reason for investing in a long-term agreement set up is to shorten procurement. This needs to be examined to identify bottlenecks in using long-term agreements.

Figure 9. Duration of DEs by Contracting Modality

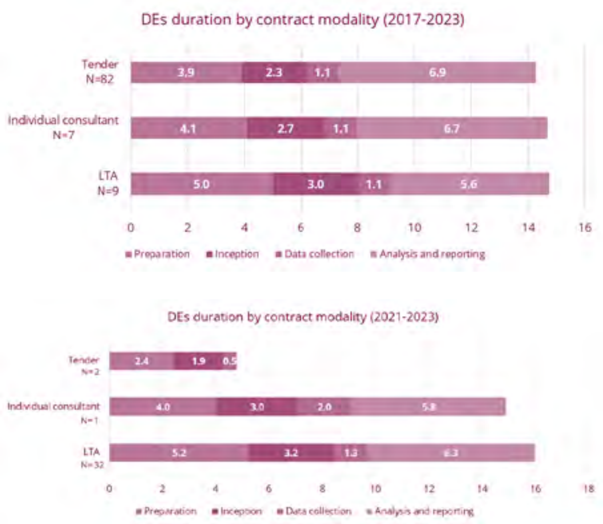


Table 4: Duration of a decentralized evaluation

Phase		Estimated/ideal	Actual median[9]
1.	Planning	As part of CSP design duration (visioning workshop to EB approval)	
2.	Preparation	9 weeks/ 2.3 months	4.1 months
3.	Inception	7 weeks / 1.8 months	2.5 months
4.	Data collection	3 weeks / 0.8 months	1.0 month
5.	Reporting	11 weeks / 2.8 months	5.9 months
	Total	30 Weeks/ 7.5 months	13.5 Months
6.	Dissemination	8 weeks/ 2 Months	-
Total (excluding planning phase)		**38 weeks / 8 months**	**15.2 months**

Duration of evaluations can differ and be affected by a range of factors, including:

- the complexity of the evaluation, including coverage and methodology;
- the level of experience of the evaluation manager in accomplishing tasks they are directly responsible for (for example, drafting the ToR);
- reassignment/turnover of the evaluation manager and country office staff, leaving gaps for periods of time and/or learning curves for new evaluation manager/country office staff as well as time needed to establish relationships with those involved;
- administrative issues with the contracting evaluators for example, multiple rounds of calls for proposals;
- turnaround times to provide inputs to the draft reports;
- efficiency of decision making processes; and
- changes in external factors (for example, sudden onset of emergency).

RB & HQ DE

In the case of **regional bureau-led decentralized evaluations**, the total decentralized evaluation duration seems to be shorter, particularly at the analysis and reporting stage. This may be explained by not having a second quality assurance layer applied (regional evaluation units do not refer to the Office of Evaluation for QA2 as indicated earlier) or the fact that these are managed by evaluation officers at the regional bureau level. However, the number of regional bureau-led decentralized evaluations is too small, so future analysis will show whether this is the trend and lessons drawn from regional bureau-led decentralized evaluations on how decentralized evaluations could be more efficient.

Joint

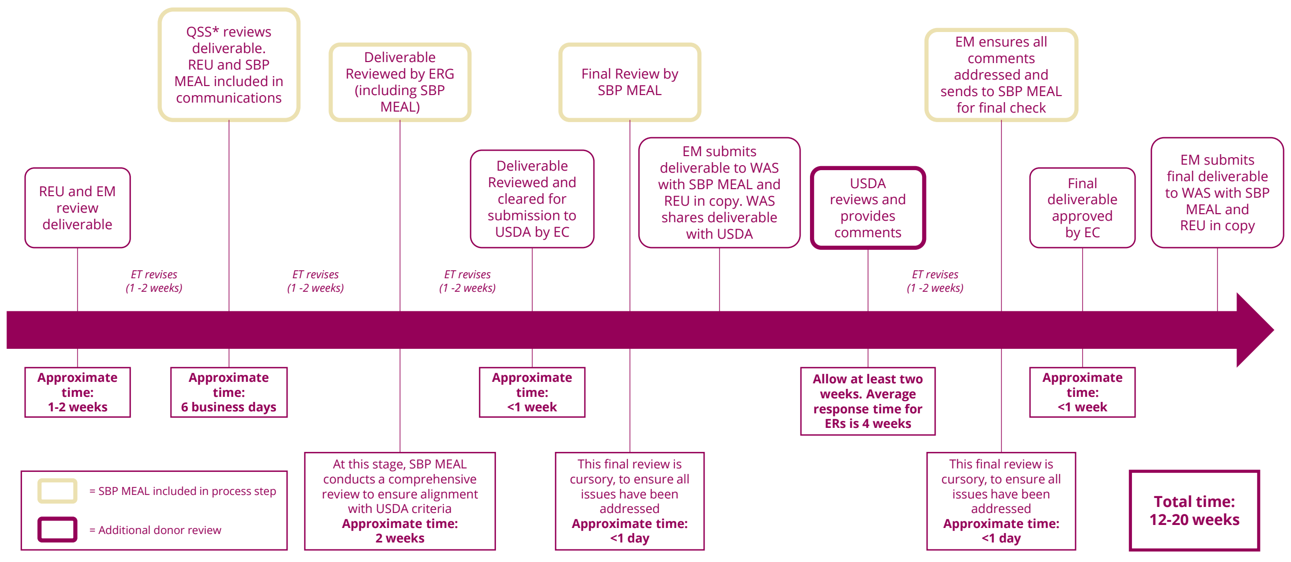
Joint evaluations tend to require lengthier consultation processes and, consequently, have a longer timeline (see the Technical Note for Joint Evaluations). The median duration of a joint evaluation from preparation to approval of the evaluation report is 17 months. Experience to date has shown that extended timelines are encountered during the analysis and reporting phase (one more month than non-joint decentralized evaluation) and the dissemination and follow-up phase (two more months than non-joint decentralized evaluations). This may be influenced by the number and type of partners involved in joint decentralized evaluations, as well as the longer approval processes required for the evaluation report and management response.

To ensure more efficient joint evaluations, it is important to jointly develop a timeline with the other commissioning agencies, communicate to all stakeholders agreed milestones and key dates and monitor progress regularly.

McGovern Dole

Evaluation of **McGovern-Dole-funded school feeding programmes** require additional time for review and approval of the ToR and evaluation report, as they need to be shared with the WFP school-based Programmes Monitoring, Evaluation and Learning (MEAL) headquarters team and USDA. Additional time is also needed for the inception reports as they must be shared with the WFP school-based programme MEAL team for review and inputs (see Figure 10).

Figure 10. McGovern Dole review timeline



*Quality Support Service, also known as DEQS

1.3. Link with monitoring plan and processes for co-led decentralized evaluations

What could go wrong?

Decentralized evaluation is not adequately planned or budgeted for in the country portfolio budget and other tools (EPB/MRE) therefore not on the radar during planning and review processes.

Even if planned, M&E team is under-capacitated to conduct the baselines/monitoring to ensure availability of quality data for use during evaluation.

Timelines for decentralized evaluation might overlap with other exercise due to timelines of projects such as MGD programmes that have different times.

Good practices

Review and advocate for evaluations to be Review and advocate for evaluations to be adequately reflected in country portfolio budget, EPB tool, annual planning process (APP), budget revisions.

Advocate, jointly as appropriate with monitoring teams for resources allocation for M&E activities.

Do not review decentralized evaluation plans in isolation, consider all M&E exercises and how best to align/complement with them.

As indicated in Section 1.2, the country office should ensure that all planned evaluations are reflected in the appropriate evidence planning and budget (EPB)/monitoring, review and evaluation (MRE) plans.[10] This should identify requirements for assessments, baselines, monitoring, review, and evaluations in a coordinated manner.

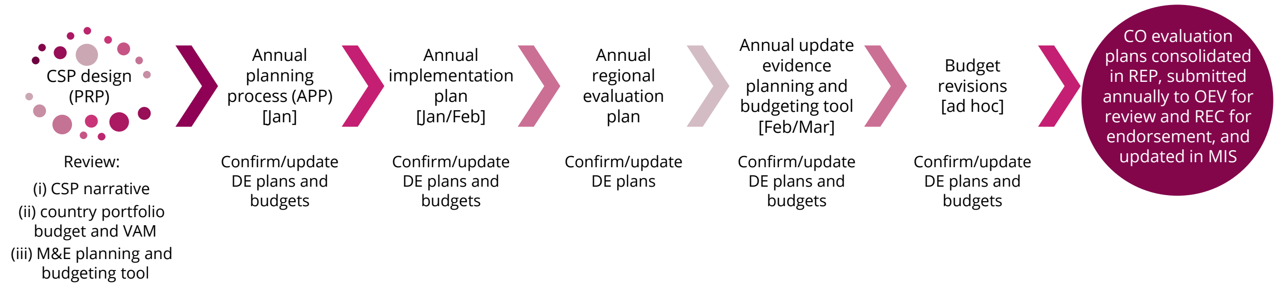
In view of the fluid and fast-moving contexts in which WFP operates, some of the assumptions made at CSP formulation stage may need to be revisited over the course of the CSP. The annual planning process and preparation of the annual implementation plan offer an opportunity to check that: the evaluation is still needed; the timing is still appropriate; the budget allocation is made and preserved; any changes in evidence requirements and their implications for planned evaluations are considered; and evaluation management responsibilities are reflected in the evaluation manager's annual PACE workplan.

Budget revisions provide another opportunity to check whether an evaluation is appropriate or whether there are other changes that need to be made on already planned evaluations to align with a change in operational context and programmatic response (for example, in case of a sudden-onset emergency). New evaluation plans can be made when developing funding proposals for a donor requiring a specific evaluation or when designing a joint programme with other partners. Before finalizing the plans, the M&E team should be clear about the intended use of the evaluation, and the availability of monitoring data and resources. The final decision rests with the director of the commissioning office, based on information prepared by the relevant programme officer, the M&E officer and the regional evaluation unit.

Until a decentralized evaluation is formally contracted, the commissioning unit may revisit its original plans and make changes including cancelling the decentralized evaluation altogether. On decentralized evaluation cancellation, see Info Brief on Reasons for cancelling and/or changing decentralized evaluations at planning and preparation stages. The main reasons for this include:

- evaluations are turned into another evidence-generation exercise such as reviews, studies etc.;
- review/ change of evidence needs by country office senior management;
- lack of capacity (human or financial) to conduct the evaluation;
- subject of the evaluation is underfunded and/or little or no implementation has occurred;
- unforeseen circumstances do not allow for the evaluation to be conducted in a timely manner;
- the overlapping timeline of another evaluation allows coverage of the evaluation subject, ensuring efficiency gains and avoiding duplication of efforts and evaluation fatigue; and
- it was expected to be a joint evaluation, but WFP does not have any decision making role, for example not represented in the evaluation steering committee.

Figure 11: Key engagement points for updating the evaluation plans.



Evaluations supported by an effective monitoring system are more likely to be credible and useful. When formulating an I/CSP, the M&E team should identify requirements for assessments, baselines, monitoring, reviews, and evaluations in line with the CSP theory of change (or logical framework in case a theory of change has not been developed). The timing of assessments, monitoring activities, reviews and evaluations should be determined for complementarity and synchronization and to ensure the most efficient use of collected data.

For most evaluation types, the start points (baseline) before an intervention[11] must be specified to allow for a 'before' and 'after' comparison. The definition of a baseline ('pre-operation exposure condition') and how to undertake a baseline study are described in the WFP Monitoring and Evaluation Guidelines and Standard Operating Procedures (SOPs) for CSP Monitoring. The standard operating procedures stipulate that a baseline be developed for all outcome indicators within three months before to three months after the start of a CSP.

Some evaluations may need data collection over and above that required for corporate monitoring. They may require a wider or deeper look at some aspects of the intervention such as partnerships, service providers, distribution sites or delivery points, etc., which cannot realistically be collected by an evaluation team during a final evaluation. This is particularly important for pilot evaluations. If so, the requirements should be specified at the design stage of the intervention. The Technical Note on evaluation approaches, methods and tools provides more information on types of data collection and explains the distinctions between primary and secondary data sources.

When identifying baseline and monitoring requirements, clear plans should be made for presenting disaggregated data (including on gender, disability and any other relevant subpopulation group).

McGovern-Dole-funded school feeding programmes include an evaluation plan, which must be approved by USDA and amended if there are any significant changes to the scope or methodology later. At least 3 percent of the McGovern Dole programme budget must be devoted to evaluation, which includes a baseline study that must be completed before any new activities can begin, a mid-term evaluation when the programme runs longer than four years, then a final evaluation.

1.4. Clarify provisions for impartiality and independence

What could go wrong?

Misunderstanding on provisions for impartiality and independence by different stakeholders.

Director of commissioning office may appoint an evaluation manager who is involved in day-to-day implementation of intervention being evaluated.

Even when an evaluation manager is not directly involved, s/he may be reporting to the person in charge of implementation raising risks to impartiality due to power dynamics.

Good practices

Proactively engage with senior management to clarify provisions for impartiality and options for management of evaluation to ensure independence and actual and perceived impartiality of the evaluation.

Share criteria for appointing an evaluation manager including during the regional evaluation committee and regional evaluation unit oversight missions to country offices as appropriate.

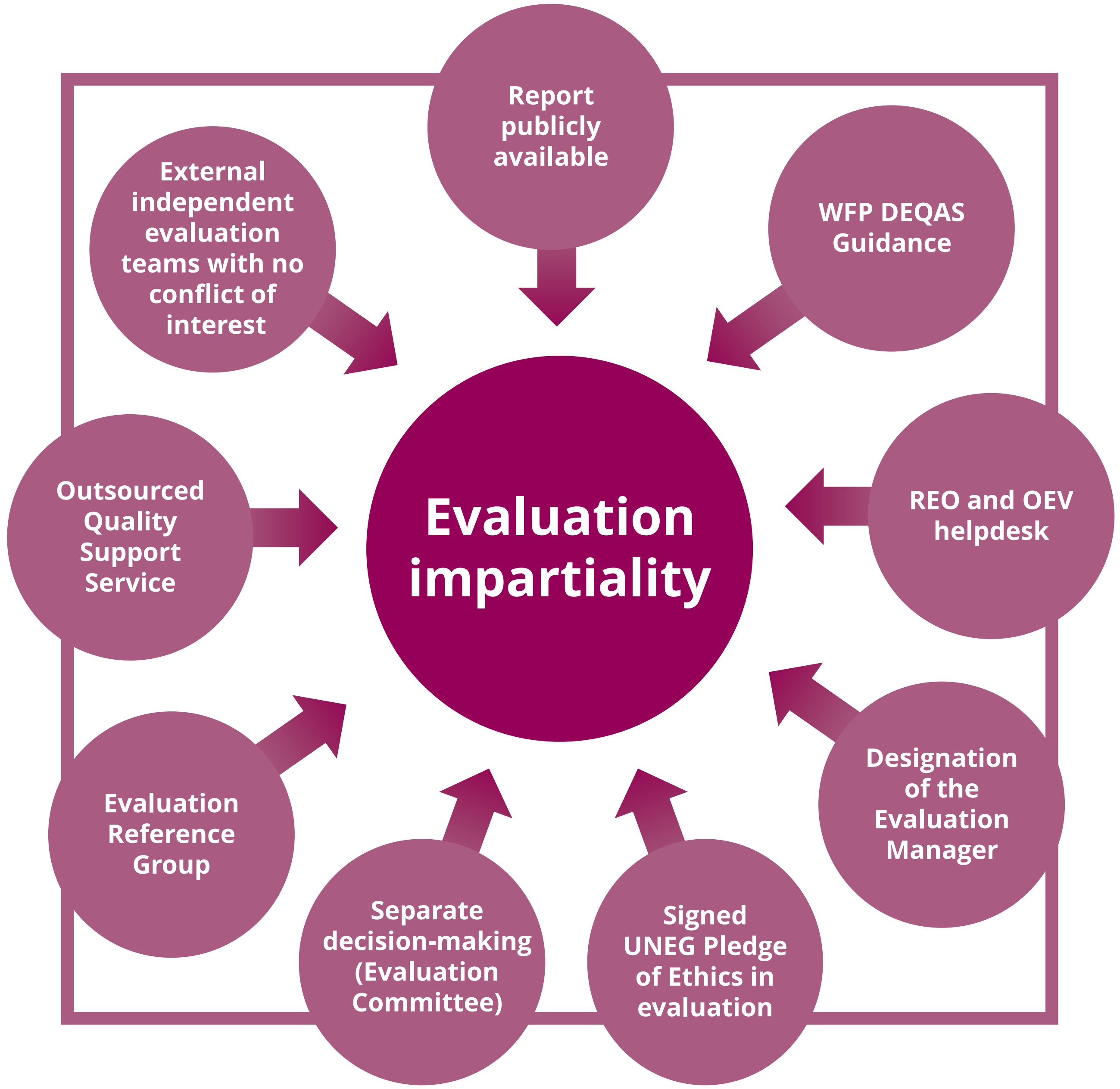
The regional evaluation unit systematically engage as individual members of the evaluation committees.

Impartiality is the absence of bias and its mitigation at all stages of the evaluation process. This is critical to the usefulness and credibility of the evaluation. The director of the commissioning office is expected to set the tone and ensure that the commissioning office staff and key partners involved in the evaluation understand the principle of impartiality and its implications. This vigilance on potential bias starts from the beginning of the evaluation process and should inform all phases (see the Technical Note on Principles, Norms and Standards).

The whole commissioning office supports the impartiality of the evaluation by engaging constructively with the evaluation team and sharing all relevant information. As custodian of data and information, the M&E team ensures that information is accessible to the evaluation team. The director/deputy director ensures that country office staff and partners understand and respect the implications of the principles of impartiality and independence.

Maintaining independence does not mean that those responsible for design and delivery of the intervention being evaluated are isolated from the evaluation process. They are key stakeholders, and their participation is essential and must be meaningful. However, they should not have any control or influence in the design or conduct of the evaluation or any veto on its findings. See Figure 12 for an overview of the provisions on safeguarding the impartiality of decentralized evaluations.

Figure 12: Impartiality provisions in decentralized evaluations



1.5. Agree on roles and responsibilities

What could go wrong?

Commissioning office takes too long to appoint the evaluation manager, PACE is closed before evaluation management is included as one of the evaluation manager's key deliverables of the year.

Management of commissioning office delegates everything to the evaluation manager and does not take on a steering role, which sends the wrong signals as to the importance of the evaluation.

Turnover of evaluation managers due to reassignments, changes in roles or consultants leaving.

Good practices

M&E team should proactively consult senior management and supervisors during PACE and annual planning processes to ensure evaluation is included as key deliverable of the year in PACE and M&E unit work plans.

As evaluation is a governance tool, include it in the management section of the annual planning process and outline the key tasks of forming the evaluation committee and evaluation reference group and chairing them.

Have co-evaluation managers/alternate evaluation managers to avoid gaps when evaluation managers leave or are absence for long periods.

The director/deputy director of the commissioning office appoints an evaluation manager in advance to allow him/her to plan ahead for the evaluation and include the evaluation management duties in his/her workplan. The choice of the evaluation manager is central to the success of the evaluation and should be made by balancing various considerations, including the key skills and capabilities outlined in Box 4. For more information, refer to WFP generic job description for monitoring and evaluation officer.

After carefully balancing the above considerations, in consultation with the regional evaluation unit (for country office-led decentralized evaluations) or the CapQual unit (for regional bureau and headquarters decentralized evaluations), the director/deputy director will appoint the evaluation manager.

RB & HQ DE

For regional bureau-led decentralized evaluations, a staff of the regional evaluation unit would be best placed to take the role of evaluation manager. For headquarters decentralized evaluations, the M&E officer of the commissioning division would be best placed to be the evaluation manager. If the division does not have an M&E officer and considering that all staff would likely be involved in the day-to-day implementation of subjects of evaluation, it is recommended the division hire a consultant to manage the evaluation.

For multi-country decentralized evaluations, the decentralized evaluation would be co-managed by one evaluation manager per country and a member of the regional evaluation unit, steered by an evaluation committee made up of at least one representative from each country office and regional bureau.

Box 4: Summary of Knowledge, skills and capabilities required for an EM

- **Strategic management**: good understanding of humanitarian and development actions within the specific country/regional context and knowledge of social-economic-political issues, United Nations reforms and national development plans and priorities as appropriate.
- **Project and process management**: solid project management skills including development of a theory of change to support design and implementation of M&E systems; and an in-depth knowledge of international norms and standards for evaluation and the ability to apply them appropriately.
- **Evaluation approaches, methodology and tools**: knowledge of different approaches, methods and tools for data collection and analysis including qualitative and quantitative methods. Ability to define key methodological requirements of an evaluation based on the key questions of interest or the problem statement of the evaluation and to critically review the evaluation design proposed by the evaluation team. Ability to validate data quality and ensure that data collection and analysis have been conducted ethically.
- **Communication and promotion of use of evidence**: ability to communicate key evidence from evaluations effectively to a range of stakeholders and create clear, influential messages to senior audiences. Active listening and conflict resolution skills are required to promote constructive engagement from stakeholders with diverse, complex and high-stake interests.
- **Ethics**: ability to demonstrate ethical behaviour and communicate WFP expectations that all stakeholders act ethically in furtherance of the ideals of the United Nations and WFP mission. Understanding of the implications of ethics in evaluation activities and ability to consistently ensure that appropriate actions are taken to safeguard ethics and escalate issues when there are violations.
- **Team and performance management**: strong interpersonal skills with demonstrated ability to manage diverse teams and solve problems effectively and sensitively. Strong organizational skills and proven ability to produce results to deadlines. Demonstrated skills in project management.

Source: WFP generic job description – M&E officer

Joint

For joint evaluations, as explained in the Technical Note on Joint Evaluations, there are three options for setting up the management of joint evaluations: a) a single evaluation manager; b) a joint evaluation management group[12] with partial representation of the commissioning units; c) a joint evaluation management group composed of all commissioning units. It is critical to discuss the available options with involved partners as early as possible.

Recognizing that the circumstances in each office vary, the director/deputy director is encouraged to seek advice from the regional evaluation unit (for country office-led decentralized evaluations) or the CapQual unit (for regional bureau and headquarters decentralized evaluations) on the advantages and disadvantages of the various options for appointing an evaluation manager outlined below.

Option 1

Country office M&E officer/regional evaluation unit team members/headquarters M&E officer

Typically, these are the best placed to take the role of evaluation manager. The advantage in the case of country office and headquarters divisions is that they know the activities under evaluation seeing that they monitor implementation, and are involved to a lesser extent in the design and implementation than programme/operations staff. This option should be feasible in very large, large and some medium country offices that have an M&E officer with the requisite capacity. In some instances, another staff member within the M&E unit may be able to take on this role if sufficiently experienced. Resource permitting, the country director/deputy country director might decide to hire a new staff member to manage the evaluation, who would be placed within the country office M&E unit. Depending on the number and complexity of the evaluation(s), this may be through a "when actually employed" contract, with the number of days established based on the required level of effort. If a country office intends to commission multiple studies in parallel, then a full-time position would be required. For the regional bureau-led decentralized evaluations, assuming division of labour within the team is by theme, the regional evaluation unit member who is focal person for the specific theme being evaluated is best placed as evaluation manager, as they will have engaged with the topic and thus be familiar with it.

If option 1 is not feasible because the country office does not have an M&E officer with adequate capacity, or because all M&E staff have been directly involved in the management and implementation of the subject being evaluated, then consider option 2.

Option 2

Country office vulnerability analysis and mapping (VAM) officer or risk management officer (regional evaluation officer (REO) for a regional bureau-led decentralized evaluation)

The VAM officer or another country office staff member such as the risk management officer, who has not been directly involved in the management or implementation of the programme/activities to be evaluated, may be appointed the evaluation manager. In the absence of capacity, the regional evaluation officer should manage the regional bureau-led decentralized evaluation himself/herself or explore the option of the risk management officer or a member of the research, assessment and monitoring (RAM) team. For headquarters decentralized evaluations, the director may appoint a staff member from a different section within the division who is not involved in the day-to-day implementation of the subject of evaluation.

If Option 2 is not feasible because the country office does not have a VAM or compliance officer or all programme officers have been involved in the management of the subject being evaluated, then the country director/deputy country director should consider Option 3.

Option 3

For country office-led decentralized evaluations, a member of the regional evaluation unit

The country office may request support from the regional evaluation unit to manage the evaluation. This should be approved by the deputy regional director. This option is to be considered when the decentralized evaluation is covering multiple countries, or in the case of single-country decentralized evaluation, for small and very small country offices. Otherwise, it would be unsustainable for a member of the regional evaluation team to manage all decentralized evaluations in the region. For headquarters decentralized evaluations, the director should consider outsourced evaluation management (Option 4c).

If option 3 is not feasible, consider one of the options under option 4.

Option 4

Three alternative sub-options are presented below:

a) Evaluation manager from another country office/regional evaluation unit: Explore the possibility of having an evaluation manager from another country office on temporary duty (TDY). For a staff member to be eligible they must have managed at least one decentralized evaluation and have completed EvalPro. The TDY could be undertaken remotely with travel to the country office at peak times of the evaluation process. An evaluation focal point within the country office team should be appointed to be coached by the evaluation manager and perform tasks that require an in-country presence.

b) Country office programme officer: A country office/regional evaluation unit programme officer could be considered if s/he oversees a different activity from that being evaluated. The advantage of the programme officer is that s/he knows the activities under evaluation and has strong linkages with the programme team generally. The selected staff member should have the ability to manage the evaluation process effectively and ensure a high degree of impartiality at all stages despite his/her involvement in the activity being evaluated.

c) Outsourced evaluation management: Part of the evaluation management role could be outsourced to an individual consultant or an evaluation firm. However, if the person is not based within the country, s/he would only be able to assume the tasks that can be performed remotely, such as drafting the ToR and quality assuring the deliverables produced by the evaluation team. This person could be based in the regional evaluation unit, and country office could cost-share the position. Other responsibilities, including consolidating the evaluation document library, preparing the field mission, or engaging with key stakeholders in-person, would need to be performed by a country office/regional evaluation unit/headquarters-division staff member. Therefore, an "evaluation focal point" within the commissioning office will need to be appointed to perform these tasks. A template for the evaluation manager's ToR is available here. This option should be of last resort when none of the previous three options are possible, as it does not contribute to enhancing evaluation management capacities within WFP.

For small country offices, the options for ensuring impartiality will be more limited as most, if not all, staff are likely to have had some involvement in the subject of the evaluation whether its design, management or monitoring and evaluation. Allocation of roles and responsibilities must take account of these levels of exposure. If a staff member has had extensive involvement in the subject of the evaluation, including: (i) playing a major role in the design team; or (ii) playing a major role in implementation, then that particular staff member should (ideally) not assume the role of evaluation manager, or be involved in sign-off of the evaluation report.

Once appointed, the evaluation manager reports to the director/deputy director on all issues related to the management of the evaluation process. In the event that the evaluation manager leaves before completion of the evaluation (for example, through reassignment), the replacement should be appointed following the same principles. The evaluation manager should engage with the regional evaluation unit (for country office-led decentralized evaluations)/the Office of Evaluation's CapQual unit (for regional bureau and headquarters decentralized evaluations) as early as possible in the process. Once appointed, the evaluation manager should sign the Pledge of Ethical Conduct and Confidentiality Agreement and then ensure that the evaluation team also signs these documents at the contracting stage.

The director/deputy director of the commissioning office sets up and clarifies the chairing of the evaluation committee. Ideally, the director would chair the evaluation committee; however, s/he may delegate this role to the deputy director (except when the deputy director is also the head of programme for country office-led and regional bureau-led decentralized evaluations). Making explicit decision on who will chair the evaluation committee allows the director/deputy director to add this to his/her workplan and anticipate future demands on his/her time. For more information about the role and composition of the evaluation committee, see the Technical Note on Evaluation Committee.

The evaluation manager should **draft the ToR for the evaluation committee** guided by the Technical Note on Evaluation Committee (which includes a template for such ToR). The evaluation committee should share the ToR with its members, informing them of its establishment. The evaluation committee's composition and the ToR should be reconfirmed and finalized during the next phase in Step 2.2.

Joint

For **joint evaluations**, a similar committee is established; although it may be called "the joint evaluation steering committee",[13] its purpose and role are the same. It should be composed of the heads or deputy heads of the agencies jointly commissioning the evaluation. The size may be larger depending on the number of partners involved as each partner may have more than one member (for example, for WFP it is important to have the director/deputy and the head of the subject being evaluated).

The **evaluation reference group (ERG)** (see Technical Note on Evaluation Reference Group) should be discussed at this early stage. The commissioning office team should reflect on the **composition of the evaluation reference group** and who will chair it, bearing in mind that steering the group requires strong leadership. The interests of key internal and external stakeholders in the evaluation should be determined, along with their potential influence of and by it. Their likely involvement in the evaluation should also be specified (see Technical Note on Stakeholder Analysis). Once key stakeholders are identified, the broad timeline for the evaluation should be discussed and shared with them to facilitate their own planning. The composition and ToR of the evaluation reference group is finalized during the next phase in Step 2.2. The regional evaluation unit should be part of the evaluation committee and evaluation reference group for country office-led decentralized evaluations.

RB & HQ DE

For **regional bureau-led decentralized evaluations**, the regional evaluation unit's evaluation manager should be part of both the evaluation committee and the evaluation reference group, and CapQual should be part of both groups for headquarters decentralized evaluations. A mix of regional bureau and country office representatives should be part of both the evaluation committee and the evaluation reference group for these evaluations depending on the focus and scope (for example, a regional bureau-led and headquarters decentralized evaluation that has country case studies would benefit from having country office staff in the evaluation reference group). For multi-country decentralized evaluations, there would be one overall common evaluation reference group with country office and regional bureau representatives. There is no need for a country office-specific evaluation reference group as this would likely lead to duplication and overburdening of country office staff.

Box 5: Purpose and role of the evaluation committee

The director of the commissioning office sets up and chairs an evaluation committee for each evaluation commissioned by his/her office. The overall purpose of the evaluation committee is to ensure a credible, transparent, impartial and quality evaluation process. This mechanism helps prevent undue influence over the key evaluation decisions. The evaluation committee is responsible for:

- supporting the evaluation manager throughout the evaluation process;
- making decisions on the evaluation budget, funds allocation, selection of the evaluation team, reviewing draft deliverables and providing feedback comments; and
- advising the chair on the approval of the final ToR, and the inception and evaluation reports.

The evaluation committee is composed of four to seven staff from the commissioning office, who, de facto, are also members of the evaluation reference group.

Box 6: Engaging with donors on evaluations.

There are four main options for donors to engage in an evaluation process:

1. Joint/United Nations system-wide evaluations: United Nations system-wide evaluations are seen as efficient accountability mechanisms, providing donors the involvement as well as results they require for reporting to their respective constituents, while maximizing collective learning. In such a case, all organizations concerned and donors would partner to prepare, resource and commission the joint evaluation.
2. As members of the evaluation reference group of a WFP-commissioned evaluation: WFP may propose to select donors that have made significant contributions to a programme to be part of the evaluation reference group. This should be decided by the evaluation committee at the start of the evaluation process and be reflected in the evaluation's ToR and in the evaluation committee's ToR.
3. Key informant of a WFP-commissioned evaluation: Evaluation teams are expected to engage with internal and external stakeholders in a systematic way, ensuring balanced and accurate findings that support relevant recommendations.
4. External evaluation: Donors may decide to commission their own external evaluation, often focusing on their broader support to a given country/region or initiative, under which WFP is one of the many stakeholders.

The Technical Note on Engaging with Donors on Evaluation provides guidance to Donor Relations Officers at country office/regional bureau/headquarters levels and country office/regional bureau staff interacting with donors on key issues related to evaluation.

Introduction

	Minimum duration	Recommended duration	Average duration
Duration	0.9 month	2.3 months	4.3 months
Key outputs	ToR, communication, learning and knowledge management plan, document library, purchase/contracts		

The preparation phase is aimed at designing the evaluation and contracting the evaluation team. Typically, the preparation phase for a decentralized evaluation should take approximately **nine weeks**. Experience to date shows that it can take longer due to different factors, such as delayed feedback from the evaluation reference group and identifying a qualified evaluator. As strong evaluators tend to be committed a long time in advance, the earlier the preparation phase is started, the higher the chance of securing a good evaluation team. **There are opportunities to reduce this by three weeks if a good first ToR is prepared and shared simultaneously with DEQS and the evaluation reference group**.

Box 7 outlines the key implications of the principles of independence, impartiality, credibility and utility at the preparation stage.

Table 5: Summary of responsibilities – preparation phase

Steps and responsibilities	CD/ DCD	CO Prog. team	CO M&E team	CO Eval Manager	Other CO units	Eval Team	RB	OEV	Other HQ units	External stakeholders
2.1 Draft the ToR (including consultations, evaluation design, timeline, budget, optional DEQS pre-ToR call)	S	P	P	L	P		P		S	S
2.2 Finalize provisions for impartiality and independence (ERG established and ToR for EC and ERG)	A	P	P	L			P			
2.3 Quality assure (using DEQS feedback and comments matrix), consult and finalize the ToR	A	P	P	L			P	S		P
2.4 Select and contract the evaluation team and finalize budget	A	P		L	P	P	S	S		
2.5 Prepare and share document library		P	P	L	S		S			
2.6 Prepare communication and knowledge management plan	S	P	P	L	P		S			

Key: A: Approve; L: Lead – takes overall responsibility to undertake the step; P: Participate actively in step; S: Support the process with specific tasks, if required

Box 7: Principles of independence, impartiality, credibility and utility

- **Independence**: The evaluation manager ensures that the independent evaluators selected have not had prior involvement in the design or implementation of the evaluation subject and have no conflict of interest.
- **Impartiality**: The evaluation manager prepares the ToR following this Process Guide and ToR template to ensure the absence of bias in terms of scope and design. An evaluation reference group, which includes key stakeholders, is formed to help steer the evaluation and reduce the risk of bias. The evaluation manager demonstrates his/her ability to maintain impartiality towards the evaluated subject, in other words, should not be biased with respect to what is being evaluated and how.
- **Credibility**: The overall evaluation design, plan and process for executing the design is reflected in the ToR, which serves as the master document for guiding the evaluation. The ToR benefits from a rigorous quality assurance process. All available data is consolidated and made accessible to the evaluators at the start of the inception phase. The selected evaluation team has a track record of producing high-quality evaluations.
- **Utility**: An explicit discussion on the intended use of the evaluation takes place with internal and external stakeholders. The evaluation questions should be focused, appropriate and relevant to users' needs and linked to the evaluation's objective(s) (accountability and/or learning). The evaluation ToR clearly sets out the evaluation's purpose, scope and intended use(s) by various stakeholders. The evaluation manager effectively steers the evaluation process to ensure adherence to planned timelines and the timely completion of the evaluation.
- **Ethics**: The UNEG ethical guidelines include a pledge of ethical conduct in evaluation to be followed by all evaluation commissioners and evaluators contracted by WFP and a checklist of ethical issues that the evaluation manager and team should consider at each phase of the process.

2.1. Draft the Terms of Reference (ToR)

What could go wrong?

The evaluation manager may not have the capacities (time and/or skills) to prepare quality ToR.

The ToR may not conform to the standard template or may not be comprehensive enough.

Staff turnover and delay nominating an evaluation manager replacement could further delay finalization of the ToR.

Limited monitoring data (or time to assess it) may limit assessment of evaluability or utilization.

For joint evaluations, a decision on which quality assurance system to use is not made on time or is not clear.

For joint evaluations, if WFP systems are not used, WFP may not have visibility of the quality assurance processes to be used.

Good practices

For country office-led/headquarters decentralized evaluations, the regional evaluation unit/CapQual should provide support in drafting the ToR, including co-drafting if needed.

Make good use of DEQS including the pre- and post-review calls with reviewers and use of DEQS feedback.

Have an alternate evaluation manager identified from the outset so they can step in in case of turnover.

If there are delays in appointment of the country office evaluation manager, the regional evaluation unit should manage the process until the country office has appointed an evaluation manager, then train him/her and handover the process. Not doing so risks the process being led by implementing staff.

Preparing the ToR is the most critical step of the evaluation process as it sets the parameters of the evaluation. It requires significant consultations and is an iterative process for which adequate time needs to be allowed. The country office evaluation manager should seek support from the regional evaluation unit, as appropriate, to ensure well-balanced consultations without overburdening stakeholders or compromising on evaluation standards. Likewise, the regional evaluation unit/headquarters decentralized evaluation managers should seek support from the Office of Evaluation's CapQual unit as appropriate. The ToR template provides guidance on ToR structure and content.

The ToR for the evaluation is the master reference document for all stakeholders.[14] It ensures firms are clear on the evaluation's requirements and deliverables when bidding for the evaluation, and ultimately, that the evaluation team and all stakeholders are clear on expectations and the process.

UNEG standards 4.3–4.7 set out key elements that a well-designed ToR should include, such as clarity on the subject, scope, objectives and sufficient rigour in methodologies to ensure a comprehensive and fair assessment (see Technical Note on Evaluation Principles, Norms and Standards).

Joint

When engaging in a **joint evaluation**, partners will need to agree on which quality assurance system to apply. As other partners may have alternative systems that are equally as rigorous as the WFP DEQAS, there is flexibility in which system to apply as long as WFP and partners are confident that the agreed system can support the delivery of a credible and useful evaluation. More than one quality assurance system should not be used in parallel to avoid duplication of work. Different agencies' quality assurance system tools should not be mixed (guidance, templates, checklists, quality support service) but rather one system should be selected and applied consistently to all evaluation deliverables.

However, useful elements of the tools (for example, using long-term agreements of another agency, guidance for handling specific issues such as engaging children, inclusion etc.) may be incorporated as and where appropriate to ensure joint evaluations benefits from agency competencies and expertise. Not all international organizations or national counterparts may be at the same stage in terms of level of maturity of their respective evaluation functions. Should commissioning partners be at an emerging stage, it would make sense to use DEQAS and for WFP to assume a lead role in evaluation management if the partners agree.

Consulting stakeholders

As part of drafting the ToR, the evaluation manager undertakes a gender-responsive stakeholder analysis guided by the Technical Note on Stakeholder Analysis, which provides further information on when and how to involve stakeholders in the evaluation process and which stakeholders to involve. This requires identifying and making early contact with the stakeholders who will form the evaluation reference group, or will be met during the evaluation fieldwork, so that they are prepared to engage sufficiently with the evaluation. It is important to ensure that stakeholders' needs and interests are considered through consultations at the start of the preparation phase.

Involving stakeholders at an early stage; keeping them informed of progress during the process; and providing opportunities for input at key stages (development of ToR, briefing sessions and reviewing draft products) will help secure support for the evaluation process and enhance the likelihood of use of evaluation findings.

Drawing from the initial stakeholder consultations that took place at the planning stage, the following points should be refined/validated to inform the draft ToR and the communication and knowledge management plan:

- What is the main purpose of the evaluation?
- What are its main objectives (balance between accountability and learning)?
- Which key questions should be addressed?
- Whose needs and interests will it serve?
- How will the evaluation be used? By whom and when?

Checking evaluation readiness (evaluability) and gathering key documents

For an intervention to be effectively evaluated (retrospective evaluation), there should be clarity on its design and logic and on the problem that it planned to address. The evaluation manager reviews the logical framework or theory of change to check that it explicitly outlines the expected changes along with planned outputs, outcomes and goals (baseline and targets). Limitations with respect to availability and reliability of monitoring data and other evidence should be systematically and transparently presented.

If the intervention does not have a clear logical framework or theory of change, the evaluation manager should consider facilitating a discussion with the programme team to map the intended results and reconstruct a theory of change from available sources. This can be conducted through a mini workshop during the preparation phase with support from the regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters decentralized evaluations). Alternatively, it can be included as a first deliverable from the evaluation team at the inception phase but will require close engagement with the programme team. More guidance is available in the Technical Note on Using Logical Models in Evaluation.[15]

The document library should be composed of the list of documents shown in Annex 1, gathered by the evaluation manager, who should also assess availability of data for the evaluation to draw on, including baseline and endline data to measure progress. The list of documents may vary depending on evaluation subject, type, etc.

Should baseline and/or follow-up monitoring data not be available, including on gender, this will require mitigation strategies including, for example, conducting additional primary data collection exercises. This should be budgeted for and implemented either as part of monitoring prior to the evaluation or as part of the evaluation. Any remaining gaps will need to be explicitly stated in the ToR. This allows the evaluation team to make a clear assessment of available data and reflect on what mitigation strategies may be needed and what the implications may be (for example, on timelines and costs).

Evaluation questions and criteria

As part of drafting the ToR, the evaluation All evaluations should be driven by questions. Evaluation questions (EQs) should be developed in line with the purpose, objectives and intended use of the evaluation. **Fewer, strategically selected, and well-defined evaluation questions are better than having too many questions that limit the depth of insights from the evaluation**. Sub-questions for each evaluation question should be limited to ensure that they are focused on answering the main evaluation question in a meaningful and in-depth way. Evaluation questions should reflect/test the assumptions underpinning the project being evaluated. It is a good practice to review evaluation questions/sub-questions to ensure that they do not overlap and there is clear delineation between them, with all of them numbered.

WFP evaluations assess WFP interventions in relation to an internationally agreed set of evaluation criteria and standards of performance. Used widely by development evaluators worldwide, the standard criteria of relevance, coherence, effectiveness, efficiency, impact and sustainability are essential components of evaluation. WFP often includes additional humanitarian evaluation criteria: appropriateness, connectedness and coverage. Depending on the subject and the evaluation questions of interest to the commissioner, a specific set of criteria might be more/less dominant in a particular evaluation. The evaluation manager should map each evaluation question or sub-question against the corresponding criteria. Which and how many evaluation criteria are addressed should be driven by the number and type of evaluation questions. The Technical Note on Evaluation Criteria and Questions provides further guidance on how to prioritize evaluation questions and link them to relevant evaluation criteria. It also includes examples of questions against each criterion.

For evaluation of **McGovern Dole-funded school feeding programmes**, at least some evaluation questions should contribute to the McGovern Dole learning agenda. This agenda will be communicated as part of the development of the evaluation plan that accompanies the McGovern Dole funding proposal and agreement. To make these donor-demanded evaluations as useful to WFP as possible, reflections on how that learning agenda may relate to and inform the WFP learning agenda would be useful, in consultation with the school-based programme team.

For **multi-country decentralized evaluations**, there should be a core set of common evaluation questions across all countries and optionally each country office may add specific evaluation questions to meet specific country needs. However, care should be taken not to have too many questions.

Considering gender equality and women's empowerment (GEWE) and disability inclusion (DI) dimensions

GEWE should be integrated in evaluation criteria and questions as per the United Nations System-Wide Action Plan (UN-SWAP) on gender equality and women's empowerment. The earlier that GEWE approaches are incorporated into the evaluation thinking, the higher the chances are that they will be thoroughly analysed and reported upon. Further information on WFP commitment to GEWE can be found in the Technical Note on Integrating Gender in WFP Evaluations, Checklist, and Quick Guide, which draw from the UNEG Guidance on Integrating Human Rights and Gender Equality in Evaluations.

Similarly, wider equity issues should be taken into consideration looking at other marginalized groups including, but not limited to, persons with disabilities (for which further guidance is available in this Technical Note on Integration of Disability Inclusion in Evaluation and the UN guidance on Integrating Disability Inclusion in Evaluations and Reporting on the United Nations Disability Inclusion Strategy (UNDIS) Entity Accountability Framework Evaluation Indicator).

The reviews of the DEQS and PHQA assessments highlighted a few key weaknesses in these areas, mainly:

- Specific information on gender, inclusion and equity, relevant/ tailored to the evaluation subject and country context, should be more explicitly included in the ToR.
- Dedicated gender and equity evaluation sub-questions should be included in the evaluation matrix, in line with the focus of overall evaluation questions.
- Relevant gender, disability inclusion and equity considerations should be noted at various points in the methodology section to ensure that it is adequately considered.
- Unanticipated effects related to GEWE and human rights should be reported upon.

Box 8: Considerations for integrating GEWE and disability inclusion dimensions in the evaluation

- The preparation for the evaluation is GEWE and disability inclusion sensitive, for example a GEWE-responsive stakeholder analysis is conducted, and GEWE concerns are integrated into evaluation questions and sub-questions.
- The ToR includes an intersectional analysis of specific social groups affected and/or spells out relevant human rights and gender equality normative instruments/policies.
- The evaluation criteria and questions facilitate the collection of GEWE and disability inclusion data.
- The communication and knowledge management plan for the evaluation considers gender dimensions and includes a GEWE-sensitive dissemination strategy.
- The evaluation team is gender-balanced, geographically and culturally diverse and has appropriate gender and broader inclusion expertise.
- Approaches, methods, tools and data analysis techniques are gender-responsive and disability inclusive (quantitative and qualitative indicators are GEWE-sensitive; data is disaggregated by sex and age and disability status where available; data collection and analysis techniques integrate GEWE and disability inclusion concerns).
- Ensure that the evaluation report reflects gender concerns, i.e. the findings, conclusions and recommendations address gender issues, where appropriate.

Selecting the appropriate evaluation design

Driven by the evaluation questions, the evaluation manager outlines the evaluation approach and methodology in the ToR guided by the Technical Note on Evaluation Approaches, Methods and Tools. This is further developed by the evaluation team at the inception stage. Most evaluations in WFP apply mixed methods (quantitative and qualitative data collection methods and analytical approaches) to account for the various dimensions of WFP interventions. The design will depend on the purpose, objectives, key questions and data available prior to the start of the evaluation.

The evaluation approach or design should bring together in an integrated and coherent manner a diverse set of evaluation methods to answer the evaluation questions selected, the scope and objectives, considering important elements such as the availability of a theory of change as well as the overall evaluation context (time, data and budget constraints). Simple evaluation questions (for example, assessing whether outputs were achieved) can have relatively simple designs, with a limited range of data collection and analysis methods. More complex questions asking WHY results were achieved or not and what the impact of that achievement was demand more elaborate evaluation methods for collecting and analysing data.

Evaluations can be summative, formative, or developmental depending on the time when they are commissioned and conducted with regards to the implementation timeline of the subject of evaluation (see Figure 13).

- **Summative evaluation**: Conducted after the programme's completion/at the end of a programme cycle mainly for accountability for results but also for learning objective for the future. It answers questions such as: "Was the programme relevant to the needs of the targeted populations? What are its outcomes? What factors affected achievement of outcomes? It is usually outcome-focused and ex-post.
- **Formative evaluation**: Used to enhance learning on ongoing programmes, helps answering questions such as: "What does and does not work? What are current strengths and weaknesses? It usually tracks early stages of a programme's implementation to learn and adjust/improve the design and implementation.

Figure 13: Evaluation approaches – When to evaluate

- **Developmental evaluation**: Used to assess developing or emerging concepts, ideas and initiatives, helps answering questions such as: "Does the programme consider external events and take limitations into account? What are general patterns across programmes?" It is focused on adaptive capacities in complex dynamic systems, in contexts that are rapidly changing, uncertain and turbulent. The Info Brief on developmental evaluation presents the main purposes and principles of developmental evaluations and provides an example of where this has been applied within WFP.

WFP decentralized evaluations use both primary and secondary data to answer evaluation questions, but usually tend to mostly rely on secondary data. Data collection typically is done through individual interviews, group interviews, observation/image recording and mini surveys. In some cases, they involve more comprehensive surveys. When access restrictions cannot be overcome, possibilities also exist to collect data remotely. **Data collection and analysis methods for each decentralized evaluation should be guided by the evaluation questions to be answered**.

When outlining the evaluation methodology in the ToR, the evaluation manager should balance between **being prescriptive enough for the evaluation team to understand expectations and being flexible enough to** leave room for the evaluation team to develop the methodology further.[16] The following questions should be considered:

- Does answering the evaluation question require representative quantitative data collection? If so, what is the population?
- Does answering the evaluation questions require qualitative data collection methods to be able to tease out/explain variations (by sex, by age, by nutrition, by region... any other criteria that matter most)?
- Should the methodology include enough qualitative data collection to discuss how and why project had results it had, in detail?
- How can the methodology outline in the ToR be made clear enough yet flexible enough so that a different team would be able to take the evaluation forward and develop it further?
- If "mixed methods" and "triangulation" are referred to, is it clear what they mean in reference to the evaluation questions?
- If the ToR proposes a sampling approach, does it allow for gathering opinions that really do vary, and that reflect experiences and perspectives of vulnerable people?

The regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters and regional bureau-led decentralized evaluations) should engage with the evaluation manager to provide technical support on developing the evaluation questions, assessing the theory of change, conducting evaluability assessment, and outlining the evaluation approach and methodology. The outline should guide the evaluation team during the inception phase to clarify if and how qualitative and quantitative methods will be mixed and at what time in the evaluation process.

DEQS advisory service at early ToR drafting stage

The DEQS is managed by the Office of Evaluation and carried out by an external firm through independent evaluation experts with English, French and Spanish language skills to provide timely feedback from an evaluation perspective.

At the preparation stage, prior to drafting the ToR, the evaluation manager can request an optional phone call with the DEQS reviewer (jointly with the regional evaluation unit for country office-led decentralized evaluations or the Office of Evaluation's CapQual unit for headquarters decentralized evaluations) to provide advice and enhance the evaluation methodology and main questions in light of the evaluation's intended use, type of questions to be addressed, data availability and resources/time available. The evaluation manager should provide a summary of the scope and intended use of the evaluation, main questions to be addressed, data availability, resources and time available to inform the discussion. To initiate this service, a call request form should be filled and submitted to the DEQS service provider, copying the Office of Evaluation's HelpDesk (wfp.decentralizedevaluation@wfp.org) and the regional evaluation unit (or the Office of Evaluation's CapQual unit focal point for the headquarters decentralized evaluation). Key options/recommendations provided during the call are then summarized by the reviewer in an email following the call. The regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters decentralized evaluations) and the evaluation manager should provide simultaneous written feedback on the DEQS service at the end of the call. See here more detailed guidance on how to access the advisory service at the early ToR stage.

Develop realistic timeline and agree a secure budget

The evaluation manager defines the evaluation timeline, including the schedule for key milestones and deadlines for each deliverable (see Template for Budget and Timeline). The timeline should be realistic, providing sufficient time for the evaluation team to develop the expected deliverables; the evaluation manager to quality assure the deliverables; and for the evaluation reference group to review and provide comments. It is advisable to plan for some contingency in case of any unforeseen challenges (for example, a sudden event delaying the field mission or withdrawal of an evaluation team member). Carefully consider a range of elements when determining the timing of the field mission (seasonality of the intervention, elections, availability of key stakeholders, etc.). The timeline should be included as an annex to the ToR.

The evaluation manager develops a detailed budget building on the overall budget estimated at the planning phase. The budget template clarifies the key determinants of budget such as: the number of days per phase; the field-based data collection activities required; and the size of the evaluation team and their daily rates. When costing decentralized evaluations, consider the following elements which may increase/reduce costs:

- What type of evaluation is envisaged?
- What scope is envisaged (in terms of the number of activities and geographical coverage)?
- What methodology do you foresee? Is additional collection of quantitative data required to answer the evaluation questions? If so, how many rounds of data collection are envisaged? What is the sample size?
- What are the costs for recruiting enumerators, travelling in the country and other logistics?
- What dissemination modalities do you foresee (see Step 2.6 below)?

Box 9 identifies factors that can affect evaluation costs.

One of the factors that may affect the cost of the evaluation and timeline is the option used for hiring evaluators (refer to Section 1.2). There are three possible options: (i) procurement through consultancy firms and research institutions that have a long-term agreement for evaluation services with WFP (here); (ii) competitive procurement process; and (iii) recruitment of individual consultants on "when actually employed" contracts. The next steps, and the schedule of the selection process to be prepared by the evaluation manager, will depend on the option selected. This would either require a vacancy announcement or an Expression of Interest and Call for Proposals, which would need to be widely disseminated. For further information, see Technical Note on Options for Contracting Evaluation Teams.

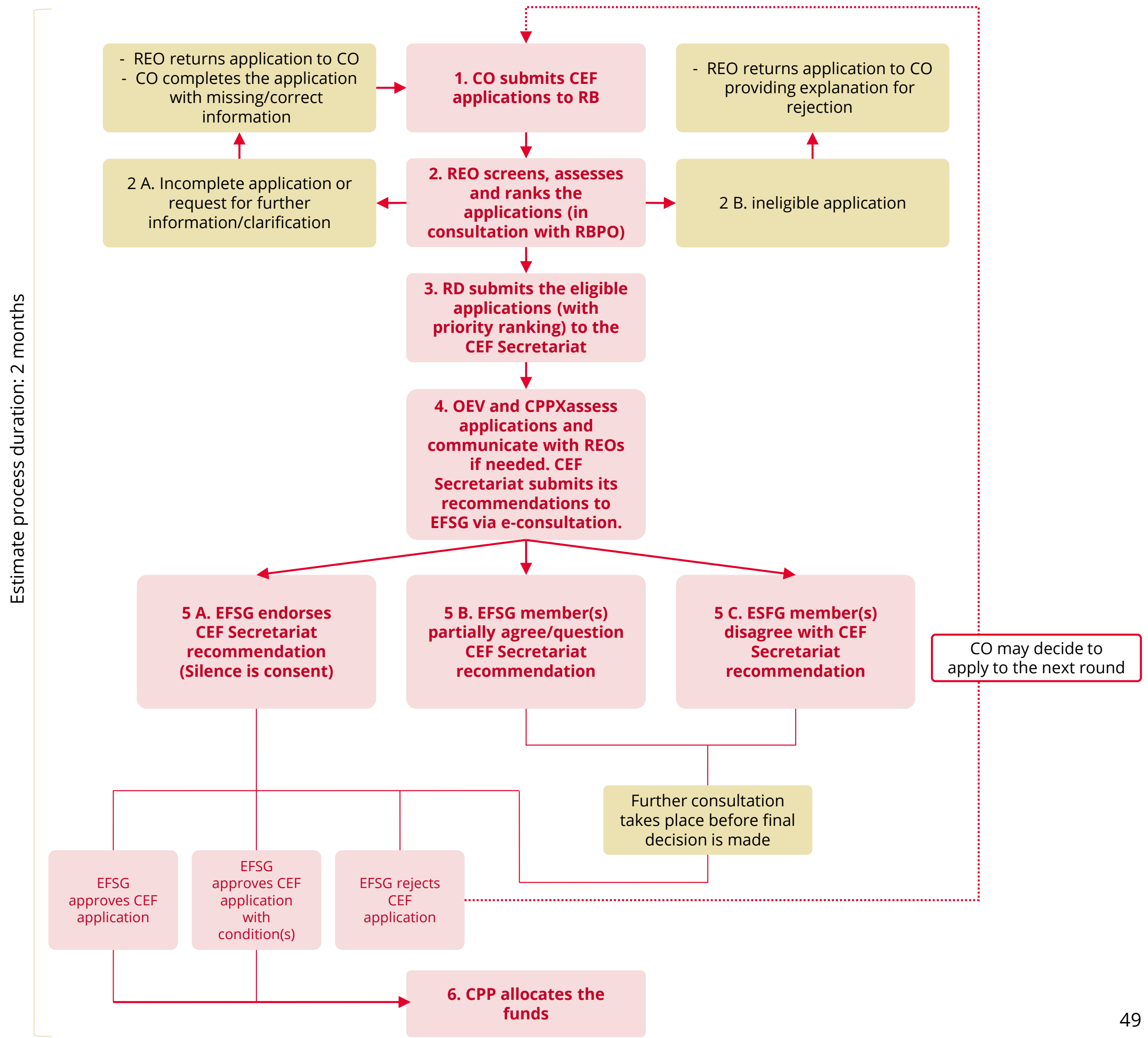
Each of the three options has implications for who makes the final decision on the team/firm selection, access to expertise, time required to recruit teams and levels of administration effort in managing the evaluation. The technical note outlines the advantages and disadvantages of each option and key considerations to be made when deciding on the contracting modality. Recruiting evaluators can take time and should be done well in advance as good consultants tend to be booked several months in advance or they tend to prefer to line up consultancies in advance.

If the evaluation was appropriately planned and budgeted but the country office is facing significant funding shortfalls, the country office can apply for the Contingency Evaluation Fund (CEF) using this CEF application form and as guided by CEF Technical Note, which describes the eligibility criteria and process for applying. A call for applications to the CEF takes place approximately every three months. If the country office intends to apply to the CEF, the evaluation manager should inform the regional evaluation unit and finalize the evaluation ToR and budget prior to submitting the CEF application form. The CEF application process takes about two months between the submission of the application by the country office to the regional bureau and the disbursement of the funds (see Figure 14).

Box 9: How to contain evaluation costs

- Ensure adequate monitoring to provide a body of data and minimize the need for and costs of primary data collection by the evaluation team.
- Link with other actors' key data collection exercises, for example, post distribution monitoring; emergency food security assessments.
- Find the right balance in the use of local/regional and international evaluation specialists.
- Limit the scope and complexity of the evaluation and limit the number of evaluation questions.
- Replace in-country inception mission with remote meetings via teleconferences.
- Conduct joint exercises where possible, sharing costs.

Figure 14: Contingency Evaluation Fund decision making process for decentralized evaluations



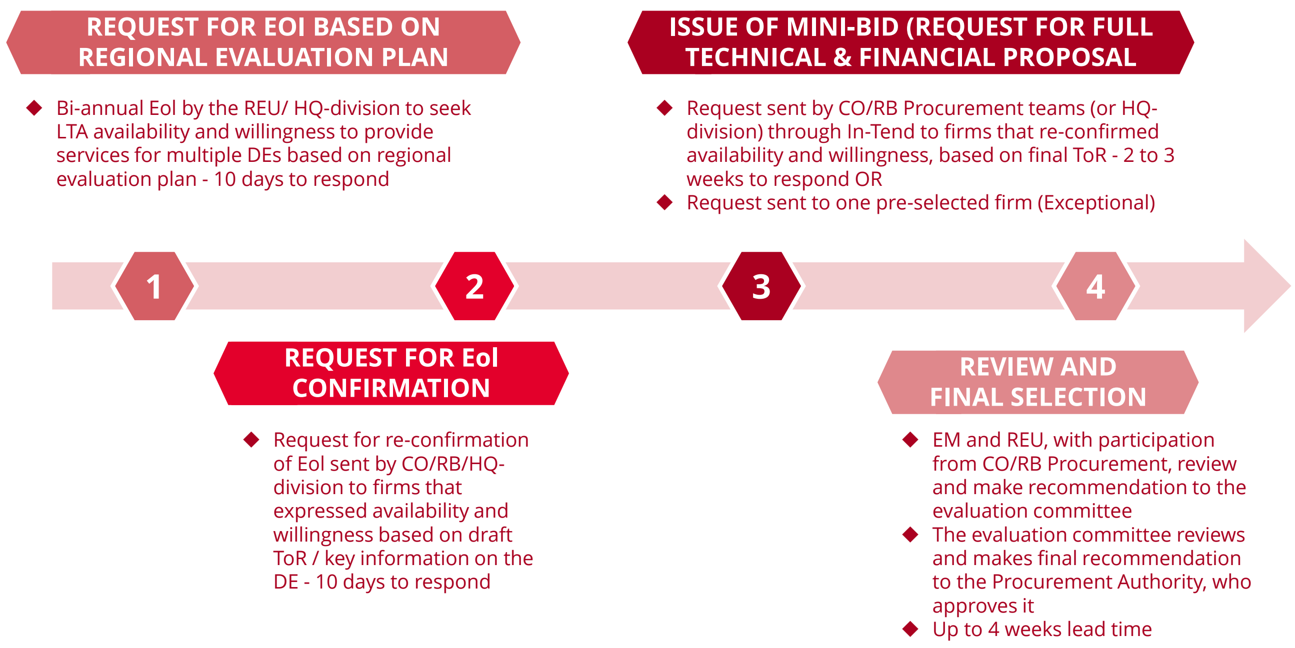
Call for expression of interest from evaluation firms

The Office of Evaluation establishes long-term agreements with consultancy firms and research institutions for the provision of evaluation services across WFP through an open, competitive procurement process. This allows successful firms and institutions to provide evaluation services using a streamlined procedure. A comprehensive list of decentralized evaluation long-term agreement holders can be found here, which includes an indicative overview of the main technical areas of work and the main methodological areas of expertise.

Given the growing demand for evaluation services, the capacity of the long-term agreement holders and availability of good evaluators can be stretched at times. Regional evaluation units should share at the start of the year a regional evaluation plan (REP) with all decentralized evaluation long-term agreement holders that qualified for their specific region, asking them to indicate their expression of interest (EoI), copying the Office of Evaluation Help-Desk as well as the evaluation manager (where one has been appointed already) and country office/regional bureau procurement units of the concerned country offices/regional bureaux, for information and inclusion in the procurement plan. Likewise, the CapQual unit should consolidate and share the planned headquarters-led decentralized evaluations. A reasonable period should be allowed for companies to respond (ten days). Firms should be updated regularly on any changes and this could be done quarterly (see an example here).[17]

The "secondary bidding" or "mini-bid" process should start at least 2-3 months before the start of the inception phase. If the country office/regional bureau/headquarters division intends to contract for the evaluation through long-term agreements, they should re-confirm with all decentralized evaluation long-term agreement holders that qualified for their specific region and expressed interest, their availability and interest to provide the required services. This should be done by the evaluation manager, in consultation with the procurement officer and the regional evaluation unit, while the evaluation ToR is still in draft form, through sharing only key information about the evaluation (scope, tentative timeline, number and profile of evaluators). The firms should be given ten days to respond. This is critical to ensure they have sufficient lead time to consider the evaluation in their schedules. The Technical Note on Options for Contracting Evaluation Teams provides further guidance. Figure 15 presents a quick overview of the process, and more details are provided in Step 2.4.

Figure 15: Overview of the mini-bidding process



2.2. Finalize provisions for impartiality and independence

What could go wrong?

The evaluation committee or evaluation reference group not set up or members not informed on time or if set up, their role and modes of engagement are not explained.

The head of subject of evaluation is the chair of evaluation committee/evaluation reference group, thus affecting perceived and actual impartiality.

The evaluation manager leaves (reassignment or other reasons) while the process is underway and there is no alternate.

Good practices

Regional evaluation unit (and CapQual for headquarters decentralized evaluations) share examples of evaluation committee and evaluation reference group ToR and supports preparation.

The evaluation manager, with support of regional evaluation unit (CapQual for headquarters decentralized evaluations) ensure ToR for the evaluation committee and the evaluation reference group is prepared and shared with members before completion of the evaluation ToR.

The evaluation manager explains to the head of evaluation subject why it is in their interest to not chair the evaluation committee/evaluation reference group but rather engage as members (using compliance reasons does not work, better to speak to people's interests).

Always keep the alternate evaluation manager (where this is identified, which is strongly advised) as well as regional evaluation unit (and CapQual in case of headquarters decentralized evaluations) in the loop as they will need to step in in the absence of the evaluation manager.

In principle the director/deputy director has already established the evaluation committee during the planning phase and assigned the chair. If the planning phase took place a long time before the preparation phase, the evaluation committee chair should re-confirm the composition and chair (see above). In any case, the evaluation manager should update the evaluation committee ToR to reflect the new membership and revised evaluation timeline.

The director/deputy director of the commissioning office establishes and chairs the evaluation reference group, which is composed of internal and external stakeholders. The evaluation manager drafts the ToR for the evaluation reference group, to be shared by the chair with its members, inviting them to be part of the group. This is an important mechanism for advising and steering the evaluation process and supporting independence and impartiality. Close attention should be paid to the selection of the members to ensure inclusive yet efficient engagement. Typically, the evaluation reference group should be composed of 8-12 members. It is important to select people who will commit to the role and add value to the process. To ensure that GEWE dimensions are integrated from the evaluation onset, the country office gender focal point and/or the regional bureau gender adviser should be members of the evaluation reference group; where regions have disability inclusion focal points/advisors and regional bureau programme advisers of relevant subjects, they should also be included in the evaluation reference group. The composition of the evaluation reference group should be finalized before the evaluation ToR is approved. For more information on the purpose and role of the evaluation reference group and how to keep it engaged, see Technical Note on Evaluation Reference Group.

2.3. Quality assure, consult and finalize the terms of reference

What could go wrong?

The ToR may have unrealistic timelines, too wide a scope and/or too many evaluation questions

Approach and methodological outline may be too vague or too contrary to what is feasible.

Delays in the evaluation reference group (stakeholder feedback) and/or inadequate time for review due to delays.

DEQS feedback doesn't add any value (is not relevant or too general).

DEQS feedback is not used thus missing opportunity to improve quality of the ToR.

Approval of final ToR takes too long, further delaying the process.

Good practices

The regional evaluation unit should provide a detailed and systematic review of the ToR and DEQS feedback to ensure it is used.

The evaluation manager proactively follows up with evaluation reference group and informs it in advance on the review timing to ensure availability and engagement. Offer to present the ToR in an evaluation reference group meeting and get timely feedback.

Request pre-review call with DEQS reviewers to discuss main contextual considerations/challenging areas that they should bear in mind when reviewing the draft ToR to ensure their feedback is relevant.

Offer to present main element of final ToR to the evaluation committee in a meeting to facilitate approval.

Once all the key elements of the ToR have been developed, the draft ToR should be systematically quality assured and finalized.

Table 5: Process map for terms of reference quality assurance and review

Evaluation manager	■ Reviews the draft ToR against the Quality checklist for ToR and corresponding decentralized evaluation ToR template to ensure that it follows the structure, includes the main content; and that there are no major contextual and/or factual errors.
REU (OEV CapQual unit for HQ DEs)	■ Reviews the draft ToR against the Quality checklist for ToR to confirm that it meets the minimum requirements for submission to the DEQS.
Evaluation manager	■ Shares the draft ToR and service request form with the DEQS (copying the Helpdesk, wfp.decentralizedevaluation@wfp.org, the REU (or CapQual unit for headquarters decentralized evaluations) and the chair of the evaluation committee).
DEQS (DEQS) reviewer	■ Acknowledges receipt and confirms return date for feedback within five working days. ■ In some instances, the DEQS project manager may decide to conduct a peer review to ensure that the assessment and feedback by the "reviewer" is appropriate. This will take a few additional days before final DEQS feedback is submitted to the evaluation manager.
Evaluation manager	■ Organizes a follow-up call with the DEQS to clarify any questions arising from the feedback received and to understand the reviewer's overall thinking on the quality of the ToR. The regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters-led decentralized evaluations) should also take part in this call. The regional evaluation unit/CapQual unit and the evaluation manager should fill out the survey on the DEQS service at the end of the call. This call is optional if the overall rating is 90 percent or above. ■ Records in the DEQS feedback form how each comment has been addressed and, for those that have not been addressed, the reasoning. ■ Finalizes the draft ToR based on the DEQS feedback and any additional guidance from the regional evaluation unit (or CapQual unit for headquarters-led decentralized evaluations). ■ Shares the draft ToR with the evaluation reference group for review and comments. When communicating with the evaluation reference group, the evaluation manager is advised to describe the previous steps taken to enhance the quality of the draft ToR (application of DEQS as well as quality assurance by the evaluation manager).[18] ■ Consolidates all comments in a comments matrix and explains how s/he has addressed them when revising the ToR. ■ Finalizes the ToR based on comments received from the evaluation reference group and submits the final ToR to the evaluation committee for approval by the chair, with offer to present in a meeting to speed up the approval.
Evaluation committee chair	■ Approves the final ToR.
Evaluation manager	■ Shares the budget and team composition with the regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters-led decentralized evaluations). ■ Shares final ToR with the evaluation reference group and with all other evaluation stakeholders.
REU (OEV CapQual unit for HQ DEs)	■ Shares final ToR with the Office of Evaluation for publication on WFP internal and external evaluation websites. ■ Updates evaluation management information system (MIS) accordingly.

2.4. Select and contract the evaluation team and finalize budget

What could go wrong?

Lengthy procurement process, delaying the evaluation process, thus reducing utility.

Difficulties in finding the right evaluators ending up with sub-par team (lacking knowledge of context, evaluation skills, subject matter expertise and/or WFP experience).

Budget submitted by selected firm is too high, beyond the available resources.

Scope in ToR is significantly wider and requires more resources than what was budgeted for at the CSP design stage/planning phase.

Good practices

Make decision on contracting option and start procurement process as early as possible.

If long-term agreement holders are used, follow the guidance to make the process as efficient as possible.

Conduct thorough review of budget submitted by firm and make suggestions for budget reductions.

Explore access to the CEF ensuring consultation with regional bureau budgeting team on unprogrammed funds.

Ensure that the team has at least one evaluator who has WFP experience and knows the context.

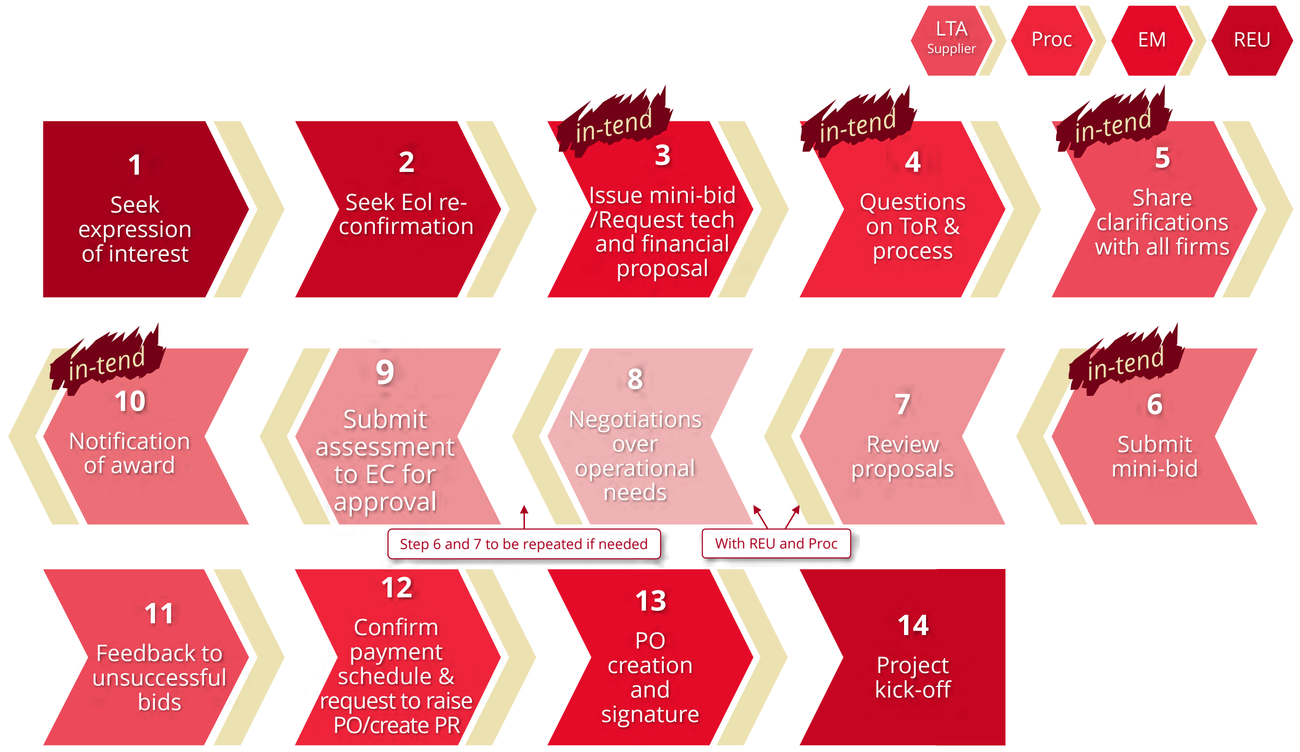
When the ToR is finalized, a request for a full technical and financial proposal should be sent to all those long-term agreement firms that confirmed their interest and availability, sharing the final evaluation ToR. A step-by-step overview is presented in Figure 16 below and further information on the process is provided in the Technical Note on Options for Contracting Evaluation Teams.

Conflict of interest: It is very important for the credibility of the evaluation that the selected evaluators are impartial and independent and that there are no conflicts of interest. Familiarity with WFP can be valuable, but the potential for bias increases when a consultant's work is solely focused on one agency. There should be no official, professional, personal or financial relationships that might cause, or lead to a perception of, bias in terms of what is evaluated, how the evaluation is designed and conducted, and the findings presented. When reviewing the curriculum vitae (CVs) of proposed evaluators, the evaluation manager should pay close attention to checking that they do not present a potential conflict of interest. When in doubt, the evaluation manager should consult with the regional evaluation unit and the Office of Evaluation. In addition, evaluators should not be recruited by the decentralized evaluation commissioning office for a period of six months after the end of the evaluation (for work that goes beyond the scope of the evaluation and the dissemination of its results). Box 10 defines what a conflict of interest entails.

Examples of conflict of interest of the evaluators include:

- past involvement in the design and implementation of WFP operations' subject of the evaluation for example, consultant supported CSP development or development of any of the CSP activities related to the evaluation; or in the development of a national policy or programme subject to a joint evaluation;
- specifically for national consultants: close association, working history, or family relationship with WFP, WFP implementing partners, competitors in the relevant country or government institution involved in the implementation of the subject of evaluation;
- having any kind of interest or ties with any WFP official or non-staff personnel of WFP;
- directly or indirectly controlling, being controlled by, or being under common control with, another related contractor or subcontractor that provides to WFP goods and/or services that are directly or indirectly related to the evaluation;
- receiving or providing direct or indirect payment from or to another related

Figure 16: Step-by-step – Main responsible stakeholder for mini-bid



Box 10: Definition of conflict of interest (extract from UNEG Ethical Guidelines)

Conflicts of interest are typically identified by a lack of independence or a lack of impartiality. These conflicts occur when a primary interest, such as the objectivity of an evaluation, could be influenced by a secondary interest, such as personal considerations or financial gains.

Evaluators are particularly exposed to potential conflicts of interest since their assessment may have major consequences on the evaluated subject, such as changes in implementation, reputation or funding. Evaluators (...) may therefore be compromised in their assessment by factors such as money or career prospects, which can lead them, for example, to provide a more positive analysis than actually deserved. It is recognized that conflicts of interest generally carry a high risk of bias and must therefore be avoided to the extent possible.

To avoid conflicts of interest, particular care should be taken to ensure that independence and impartiality are maintained. For example, evaluators should not evaluate subjects for which they have worked or had responsibility in the recent past, or in which they have been financially involved. Similarly, they should not evaluate any organization, department or team to which they are presently applying or where there is a significant possibility of being hired in the near future. Conflicts of interest, however, may go beyond such basic rules. They can be manifold and at times ambiguous. It is therefore all the more important to stay aware of any apparent or potentially arising conflicts of interest and to respond to them adequately.

contractor or subcontractor that provides to WFP goods and/or services that are directly or indirectly related to the evaluation;

- participating in another related contract either individually or as a partner in a joint venture or any similar partnership; and
- having participated, directly or indirectly, in the preparation of the specifications, scope of work or terms of reference included in the solicitation documents provided by WFP.

Evaluator skill set: It is also very important for the credibility of the evaluation that the evaluation team has the required expertise and experience to conduct the evaluation. For this, you can check the UNEG Evaluation Competency Framework. Skills required will vary according to the type of evaluation but will cluster around the core skills shown in Box 11. The evaluation manager is responsible for identifying and pre-selecting the evaluation team and checking combined competencies against the evaluation's skills requirements as per the ToR.

Box 11: Skill set for evaluators (adapted from UNEG Evaluation Competency Framework)

The skills and qualifications needed by the evaluation will vary depending on the sectors, type of evaluation and implementation arrangements; however, the following are important considerations:

- **Ethics and integrity**: The evaluator has knowledge of UNEG Ethical Guidelines, is able to understand their implications for the evaluation process and communicates this to others; is able to ensure the protection and confidentiality of evaluation subjects, including that the evaluation will not place the individual/organization at risk.
- **Evaluation expertise**: The evaluator's skillset includes the methodological skills to design and implement evaluations and knowledge of ethical standards in evaluations.
- **Thematic/sectoral expertise**: The requirement for different thematic/sectoral expertise will vary from evaluation to evaluation (though GEWE and broader inclusion-related skills should always be mainstreamed) and should be clarified in the ToR.
- **Knowledge of the context**: The evaluator has a sound understanding of country/regional context contributes to the soundness of the analysis.
- **Knowledge of WFP**: Having knowledge of WFP (and willingness to challenge the organization on the basis of evidence) means that team members can swiftly immerse themselves in the specific subject of the evaluation, rather than learning about the organization. This should not be required for all team members.
- **Language skills**: The evaluator has the ability to communicate in the official language of the country.
- **Interpersonal and communication skills**: The evaluator has the ability to communicate and liaise with a range of stakeholders in a way that enhances the credibility of the evaluation process. This also includes communication skills and the ability to facilitate group discussions and present findings in an accessible way.

Box 12: Questions to consider when selecting an evaluation team

Is the team leader sufficiently experienced with a demonstrated ability to manage a team?

Are the evaluators credible in terms of competence, reputation and a strong track record?

Does the evaluation team have the necessary mix/complementarity of skills, including evaluation skills, experience and country knowledge for this evaluation?

Do the evaluators have appropriate interpersonal skills?

Does the team have the ability to remain balanced, impartial, objective and constructive?

Has the team worked together previously? Will they function effectively as a team?

Is the evaluation team gender and diversity balanced with members from the countries or regions concerned?

When selecting an evaluation team directly as individual consultants, the evaluation manager is responsible for organizing interviews with all evaluation team members and checking references with credible referees based on the consultants' latest assignments, outlined in their CVs. If the consultant has undertaken previous evaluations for WFP, the evaluation manager in charge of the evaluation should be contacted. The regional evaluation unit and the Office of Evaluation can help to identify him/her.

If the evaluation team is contracted through a long-term agreement firm, to facilitate an objective assessment of the proposals, the evaluation manager should use the template for the assessment of proposed evaluation teams (scoring grid), in consultation with the regional evaluation unit for country office decentralized evaluations and with CapQual for headquarters-led decentralized evaluations. There should be a consensus on the final assessment.

Negotiations over operational needs

Once all the proposals have been technically and financially assessed and scored, negotiations over operational needs may take place with the preferred bidder(s) (that is, those bids that scored the highest after the technical and financial assessment of the mini-bids; this can be one or more bid(s)).

The scoring grid should record any clarifications needed and requests for adjustment to be asked to the preferred bid(s). The preferred bidder(s) shall be informed about the elements that have led to the negotiation being necessary. These negotiations over operational needs can include:

- adjustments in the number of days to ensure adequate time is spent in developing the tools, in the field and in the analysis while remaining within the planned budget;
- adjustments in the size of the team to ensure it is allows coverage of the full evaluation scope without being too big, but no negotiation of the daily rates;
- clarifications and adjustments of the technical proposal to ensure a good understanding of the assignment and a sound evaluation approach which adequately responds to the ToR; and
- changes in the team member's profiles based on, first, a reference check; and second, an interview with the selected team members/ team leader when there is no clarity on the adequacy of a profile based on the review of the CVs and when the reference checks are not conclusive

Reference checks should be done internally if the team member has WFP evaluation experience; and externally if the team member doesn't have WFP evaluation experience, through three references checks (when possible) to avoid subjectivity. These can be shared by the long-term agreement firm, as that firm is expected to have conducted these background checks and interviews prior to submitting the proposal.

Interviews for decentralized evaluations will be led by the evaluation manager, with the participation of the regional evaluation unit and, when available, the country office procurement officer. For regional bureau-led decentralized evaluations, the regional bureau procurement officer should participate, when available. For headquarters decentralized evaluations, the Office of Evaluation's CapQual will participate in the interviews.

The improved offer(s) should be resubmitted by the preferred bidder(s) and reassessed following the same procedures.

The selected proposal based on the scoring results of the secondary bidding is presented for review to the evaluation committee, which sends the final recommendation to the procurement authority. Once the composition of the evaluation team and budget are approved by the procurement authority, the next steps will vary depending on the contracting option. See the Technical Note on Options for Contracting Evaluation Teams for more information.

Regardless of the selected option, all evaluators must sign and submit the Pledge of Ethical Conduct in Evaluation and the Confidentiality Agreement to the evaluation manager before their contract is finalized. The evaluation manager is responsible for collecting and filing the signed copies along with all other contract documentation.

Joint

For **joint evaluations**, it is essential to build consensus on the evaluation team selection among the commissioning agencies to avoid challenges at a later stage. The process of selection should include the core members of the evaluation committee from commissioning partners at each stage, ensuring communication of decision to all members including those that may miss meetings to avoid any partner feeling excluded from decision making.

2.5. Prepare and share a document library

What could go wrong?

Insufficient research and available materials/reference/documents.

Too many documents collected and poorly organized leading to overwhelmed evaluators who may not know which ones are critical and which ones are not.

Links in some documents may be only accessible internally.

Evaluators not knowing how valid/credible the shared data is.

Good practices

Establish a solid library in advance and use a checklist to indicate all documents and data to be shared.

Structure the library in a meaningful way, and signal to the evaluators the critical documents that they should prioritize.

Note to the evaluators any challenges or concerns regarding any of the data/documents shared

Once the team is selected, it should be provided with a comprehensive document library. Providing full access to WFP information, documents and data is essential to guarantee the impartiality and credibility of the evaluation (see WFP Directive on Data Disclosure CP2010/001). The document library lays the foundation for the evaluation team in terms of its initial preparation and understanding of the evaluation subject. Having a comprehensive and well-organized library will also accelerate the inception phase and minimize evidence gaps. Be mindful of the time needed by the evaluation team to review these documents and only select relevant information/documents.

Building on the initial consolidation of documents made ahead of drafting the ToR, the evaluation manager, with support from relevant country office/regional bureau/headquarters staff, is responsible for ensuring that all relevant data sources are available to the evaluation team in an organized manner. The evaluation manager could use a checklist to indicate all documents and data to be shared with the evaluation team.

2.6. Prepare communication, learning and knowledge management plan

What could go wrong?

Some stakeholders may be left out of the plan.

The plan may be omitted/ not prepared at all or prepared very late in the process.

The plan only includes obvious communication channels/products (for example, sharing reports by email) thus limiting opportunities.

Good practices

Include a draft plan in the ToR to allow stakeholders to provide feedback when reviewing the ToR draft.

The regional evaluation unit (CapQual for headquarters decentralized evaluations) should review and provide inputs on the plan well in advance.

Ensure draft plan is costed and included in the evaluation budget.

An initial communication and knowledge management plan should be developed during the ToR drafting stage and annexed to the ToR. This will be refined and adapted throughout the evaluation, with the support of the programme team, reporting officer and partnership officer. It is a crucial step in ensuring adequate engagement during the process, follow-up and knowledge building (see the Template on communication and knowledge management plan and the Technical Note on stakeholder analysis). The plan should make clear the evaluation team and commissioning office respective roles and responsibilities, timing and means of communication. It helps ensure that evaluation results are widely disseminated among all relevant stakeholders. Relevant internal and external audiences should be identified and both content and communication approaches should be tailored to the respective audiences. Any translation needs should also be identified. The costs for dissemination products and translation should be included in the overall budget for the evaluation.

Joint

For **joint evaluations**, opportunities for wide dissemination can be explored by pooling resources and expertise; by expanding WFP audiences with those of the other agencies; and by using partners' contacts and communication channels. By systematically posting evaluation products in the respective websites of partners as they are approved, the evaluation may have a wider reach.

Introduction

	Minimum duration	Recommended duration	Average duration
Duration	0.7 month	1.8 months	2.1 months
Key outputs	Inception report (data collection tools, field work plan, evaluation matrix)		

The inception phase requires that the ToR is final; the evaluation team is recruited; and the document library is in place. Typically, an inception phase for a decentralized evaluation should take approximately **seven weeks**. Experience to date shows that it can take longer due to different factors such as the complexity of the subject of the evaluation and delays experienced while seeking feedback from stakeholders. **There are opportunities to reduce this by three weeks if a good first inception report draft is received and shared simultaneously with DEQS and the evaluation reference group**.

The inception phase serves to ensure that the evaluation team develops an in-depth understanding of the subject of the evaluation and other requirements indicated in the ToR and can translate them into a workplan according to which the evaluation will be carried out. The workplan includes the process of the evaluation as well as the data collection and analysis methods selected to answer the evaluation questions.

Table 6: Summary of responsibilities – inception phase

Steps and responsibilities	CD/ DCD	CO Prog. team	CO M&E team	CO Eval Manager	Other CO units	Eval Team	RB	OEV	Other HQ units	External stakeholders
3.1 Conduct evaluation team orientation		P	P	L	P	P	P			
3.2 Undertake desk review of documents		S	S	S		L				
3.3 Conduct inception meetings	P	P	P	S		L	P			P
3.4 Prepare the draft inception report				S		L				
3.5 Quality assure the draft inception report				L		P	P	S		
3.6 Finalize and approve inception report	A	P	P	L	P	P	S			P

Key: A: Approve; L: Lead – takes overall responsibility to undertake the step; P: Participate actively in step; S: Support the process with specific tasks, if required

Box 13: Principles of independence, impartiality, credibility and utility

The principles of independence, impartiality, credibility and utility, as well as ethics, have a number of specific implications at the inception phase as follows:

- **Independence**: To ensure the adequate implementation of evaluation standards and principles, the evaluation manager and stakeholders should provide the evaluation team with access to key informants and all available data at the start of the inception phase.
- **Credibility**: The evaluation methodology should be appropriate to respond to the evaluation questions. The methods for data collection, analysis and interpretation should be transparently documented in the inception report. The sites for field visits and stakeholders consulted should be selected according to explicit criteria and the sampling strategy should be detailed. The inception report should benefit from inputs from key stakeholders and a rigorous quality assurance process.
- **Utility**: The evaluation team, with the support of the evaluation manager, should ensure an efficient evaluation process as per the timeline to avoid late completion of the evaluation.
- **Ethics**: Evaluators should behave ethically in all interactions with stakeholders. Adequate ethical safeguards should be indicated in the inception report, and an assessment of risks and mitigation measures included. Evaluators should exercise commitments to avoid conflicts of interest and discretely report any discovery of apparent misconduct (in line with UNEG Ethical Guidelines).

The inception phase involves initial analyses of background materials from the document library and discussions with stakeholders (in person or virtual) to provide the evaluation team with a greater understanding of issues and concerns related to the subject of the evaluation. The evaluation team is also expected to analyse critically and further expand upon some elements of the ToR, notably ensuring that the evaluation subject, context and scope are correct, relevant, up-to-date, appropriately nuanced and politically sensitive.

The inception report clarifies the whole approach to the evaluation and, together with the ToR, constitutes the main reference tool for guiding the evaluation and checking its progress. The inception report helps establish clarity and mutual understanding across parties during the evaluation and is meant to clearly articulate what the evaluation is focused on; how the work is to be performed; who does what; what is to be produced; and when deliverables are expected.

Within this framework, the main objectives of the inception phase are to:

- critically review or (re)construct the theory of change of the subject of the evaluation;
- deepen the evaluability assessment that will have been conducted during preparation phase;
- fine-tune evaluation scope, and questions/ sub-questions, also in view of the evaluability assessment;
- develop the evaluation methodology, considering ethical issues, risks and mitigation strategies;
- develop a detailed evaluation matrix structured around the evaluation questions, sub-questions and lines of inquiry for each;
- develop and test data collection tools/ protocols;
- interview key informants to inform the evaluation design;
- confirm criteria and selection of field visits;
- deepen and finalize stakeholder mapping and analysis;
- refine the communication and knowledge management plan;
- develop a detailed workplan with roles and responsibilities for the team and deadlines for deliverables;
- elaborate a field visit schedule; and
- enhance the knowledge of the evaluation team on WFP programming, processes and data systems, as relevant for the decentralized evaluation.

3.1. Conduct evaluation team orientation

What could go wrong?

Insufficient time is allocated to orientation meetings with the team to clarify expectations and requirements, especially if evaluators are new to WFP evaluations.

Evaluation team has negative team dynamics and struggles to work together.

Evaluators are not tactful/diplomatic in asking sensitive questions leading to complaints.[19]

Good practices

Review the agenda to ensure key staff are invited to and are available for orientation meetings.

Include specific sessions on admin issues and to discuss senior management expectations with the evaluation team.

Engage closely with the evaluation team to obverse any issues between members and resolve them early.

The orientation meeting should take place prior to the inception mission (if one is planned), either in person or remotely. The evaluation manager, who is responsible for organizing it, should ensure the participation of the evaluation team, the evaluation committee chair, the programme team, and the M&E team.

The orientation meeting's key purpose is to present the ToR and expectations and should include:

- key features of the subject of the evaluation;
- the main users of the evaluation and its purpose/intended use;
- relevant data sources (corporate policies and strategies; monitoring data; previous evaluations/research; and data sources on the context);
- gaps or limitations in the data and priority areas of analysis/data collection to fill such gaps;
- the expected deliverables, timeline and deadlines;
- mutual expectations - WFP and the evaluation team, but within the evaluation team, who may not have worked together previously (for example on report writing by individual team members);
- the review of DEQAS quality standards, inception and evaluation report templates, quality checklists and arrangements for quality assurance (including the role of the team leader, the firm quality assurance adviser, the WFP evaluation manager, the regional evaluation unit and DEQS), and PHQA of the final evaluation report, as well as clarifying the process for ensuring the evaluators address stakeholder comments through rounds of commenting processes;
- confirmation that the evaluation team is aware of its obligations in relation to all evaluation activities in line with standards set out in UNEG guidance, in particular the Pledge of Ethical Conduct and Confidentiality Agreement (confirm that all team members have signed these documents and discuss key implications);
- the roles and responsibilities of all key stakeholders involved in the evaluation;
- the communication modalities and rules of engagement, including:
  a. To whom, how often and how the evaluation team will report to WFP
  b. How the evaluation team is expected to engage with external stakeholders
  c. To whom, when and how the evaluation results will be disseminated
  d. **Reporting of wrongdoing**: Should the evaluators uncover allegations of wrongdoing and misconduct in the implementation of a programme either by a WFP staff member or partner (including fraud, food diversions, misuse of WFP assets, harassment, sexual harassment, etc.), the evaluation team leader should report the allegations to the WFP Office of Inspection and Investigation (OIGI) through the WFP hotline (+39.06.6513.3663; investigationsline@wfp.org). The team leader should inform the evaluation manager, country office management, the regional evaluation unit and the Director of Evaluation that there are allegations of wrongdoing and misconduct, without breaking confidentiality
  e. **Conflict resolution modalities**: Should there be conflicts within the evaluation team they should be discussed with the firm and the evaluation manager. Should there be a disagreement with the evaluation manager, this should be brought to the attention of the evaluation committee chair and the regional evaluation unit. If the issue remains unresolved, the Office of Evaluation (senior evaluation officer in charge of CapQual unit) should be approached for guidance; and
  f. next steps, including scheduling inception meetings/teleconferences.

3.2. Undertake desk review of documents

What could go wrong?

Evaluators start work before expectations are clarified.

Key documents and relevant data may not be available to the evaluators on time.

Information overload for evaluators (too many documents versus the limited time to prepare the inception report) with no signal on what they should prioritise.

Evaluators miss reviewing key documents before engaging stakeholders.

Evaluators do not critically examine validity/correctness of secondary data provided.

Good practices

Systematically highlight the important documents that the evaluators should pay attention to.

Signal to the team any known limitations with secondary data, including gaps.

Provide available data analysis (where possible the raw data as well in case evaluators need to re-analyse).

It is very important that all evaluation team members are familiar with the main documents so that best use can be made of the inception mission and fieldwork time. The evaluation team is expected to conduct a desk review and analysis of the documents stored in the document library prior to the inception mission and particularly those signalled as priority by the evaluation manager.

Analysis of key design/implementation documents and the logical framework/theory of change (or alternative sources that set out the intervention logic) is particularly important; along with identification and analysis of the assumptions made (see Technical Note on Using Logical Models). From this analysis, implications for the evaluation can be established. This phase should also include the review of context-specific documentation (such as policies, strategies, assessment reports and government programmes) if applicable.

While the document library should provide most of the key documents, the evaluation team is likely to require additional documents relevant to the subject of the evaluation during the evaluation to fill evidence gaps and triangulate information. Any additional information required should be provided by the evaluation manager with the support of the M&E and programme teams.

3.3. Conduct inception meetings

What could go wrong?

Non-availability of key stakeholders for inception meetings.

Evaluators have not read basic documents and they ask very basic questions during the meetings, which frustrates stakeholders who may have spent time sharing the documents.

Evaluators ask evaluative questions thus duplicating questions to be asked in the next phase (stakeholders get frustrated when they have to repeat answers to same questions).

Good practices

The evaluation manager ensures stakeholders are given advance notice about the meetings and next steps.

Use programme meetings and other forums to provide updates so that internal stakeholders can support with reaching out to relevant external stakeholders.

Active involvement of leadership of the commissioning office sends the correct signals to others.

Actively remind evaluators of the importance of reading the relevant documents prior to inception meetings and of staying focused during inception.

The evaluation manager should organize a series of inception meetings for the evaluation team to meet with key internal and external stakeholders. S/he may participate in these meetings, provided that the team leader agrees. Should the evaluation manager be part of some inception meetings, it would be useful if the team leader clarifies at the start of the meeting that "this discussion is taking place during the inception phase; however, once the evaluation enters the data collection phase, WFP staff will not be participating in meetings to safeguard the principle of independence".

Inception meetings may take place with:

- the evaluation committee chair;
- WFP staff directly involved in implementation of the evaluation subject;
- the evaluation committee or evaluation reference group members;
- regional bureau staff;
- headquarters staff from relevant divisions; and
- external stakeholders who have an interest in the evaluation - for example, government, United Nations agencies, donor or cooperating partners.

The regional evaluation unit (or CapQual unit for headquarters-led decentralized evaluations) can also provide useful support to the evaluation manager by participating in key meetings with the evaluation team, clarifying expectations and ensuring stakeholders are given enough advance notice about the meetings and next steps.

The evaluation manager in consultation with the team can decide to either conduct inception meetings remotely or face-to-face. Holding face-to-face meetings has cost implications when the team leader is based in another country but can be more effective in achieving the intended results of the inception phase (see Box 15). When inception missions are planned, this should be explicitly stated in the ToR with associated costs included in the budget.

If an inception mission takes place, the evaluation committee and/or the evaluation reference group should: (i) provide a briefing at the start of the mission; and (ii) be debriefed by the team at the end of the mission. The inception mission will usually last five working days.

Box 14: Purpose and content of inception meetings

- Clarify the ToR and the context in which WFP operates and the subject that is being implemented.
- Discuss key topics of interest with the commissioning office and stakeholders.
- Deepen stakeholder analysis and discuss how different stakeholder perspectives will be included.
- Discuss evaluation matrix and evaluation questions (see Technical Note on Evaluation Matrix and Technical Note on Evaluation Criteria and Questions).
- Determine viability of sampling criteria and strategies, and sites based on them.
- Collect additional documentation/data (monitoring reports, data sets etc.); identify any gaps.
- Identify data limitations and how these will be addressed, any implications.
- Agree on administrative and logistics arrangements for the field mission (such as travel to field sites, translation/interpretation, modes of transport etc.).
- Take account of access/security/protection considerations.
- Pilot and refine data collection methods and tools as appropriate.
- Discuss specificities of the interventions being evaluated including monitoring and reporting.

3.4. Prepare the draft inception report

What could go wrong?

Evaluators do not use WFP template for inception reports and/or do not follow the guidance.[20]

Poor quality inception reports can be due to many reasons (for example, evaluators take other assignments resulting in less time allocated to the evaluation).

Proposed methodology in the inception report deviates significantly from the ToR and/or has additional costs implications beyond what is available.

Some methods suggested without corresponding data collection/analysis plan (for example, observations without observation guide; surveys planned but no pre-testing of tools).

Significant delays in submitting the draft inception report.

Good practices

Ongoing discussion/engagement with the evaluation team members to ensure they understand requirements.

If needed, raise issues with the team leader and/or the firm early when issues are detected.

Focus on review of data collection tools when reviewing draft inception report; do not approve without all annexes.

The evaluation team leader prepares the draft inception report with the support of the evaluation team, which ensures the evaluation team's ownership of the process. It remains internal to WFP and is shared with the evaluation reference group but is only published on the WFP internal website.

The inception report should follow the structure and include the content provided in the template for Inception Report and should meet DEQAS quality standards (see Quality Checklist for Inception Report). To generate trust in the credibility of the evaluation, as well as its independence and impartiality, the inception report should be:

- well-informed by analysis of the documentation provided and the inception meetings held;
- technically sound with a robust methodology (Technical Note on Evaluation Approaches, Methods and tools) based on reliable and triangulated data including the perspectives of a diversity of stakeholders;
- feasible and realistic based on inception discussions on resources, access, security, data availability, etc.;
- complete, with data collection tools corresponding to all proposed data collection methods; and
- complete with analysis plan corresponding to all proposed data analysis methods.

The evaluation matrix is the most critical part of the inception report. It breaks down the main evaluation questions from the ToR into key sub-questions and includes indicators, methods and data sources. The matrix sets out how the evaluation's methodology will be operationalized by the evaluation team. More guidance is available from the Technical Note on Evaluation Matrix.

The evaluation team should identify approaches to personal data and privacy including on which data needs to be protected, who is accountable for data protection, how should personal data be protected, and data retention plan.

3.5. Quality assure the draft inception report

What could go wrong?

Incomplete/poor-quality draft inception report submitted by the evaluation team, missing some elements.

Delays in submitting draft inception report thus putting time pressure to review processes.

DEQS feedback on draft inception report not useful and/or useful but not utilized to improve the draft.

Methodology in the inception report significantly deviates from the one proposed in the technical proposal of the firms upon which contract was issued.

Good practices

Ongoing discussion/engagement with the Do not submit incomplete inception report draft to DEQS, instead return to evaluation team and request they complete/revise and remind them to check it against the inception report quality check list.

Highlight to DEQS key challenging areas for review and feedback.

Ensure continuous discussion and engagement with the evaluation team, reminding them of expectations.

Discuss DEQS feedback during the call with the reviewer and remind the team that they are expected to update the matrix on how they have used the feedback and/or rationale.

Box 15: Suggestions to ease the inception report review process

- **Context**: Are there any major factual errors to bring to the team's attention?
- **Site selection**: The selection of field site visits has been made to allow a broad representativeness of WFP operational areas. Do you have any reservations about the selection as is presented in the draft report? If so, why? What adjustments do you suggest? Has the team missed any main areas/sites?
- **Mission schedule**: Has the team prioritized external stakeholders to be interviewed based on available information? Does the sampling frame include diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? Have any key stakeholders been missed? If so, which should be included? Are the assumptions for travel time from one location to another realistic? The country office may make suggestions for grouping site visits or interviews in different ways to contribute to efficiencies, noting that the team leader will make the final decision.
- **Evaluation matrix**: Are there any evaluation questions as per the ToR that are not addressed in the evaluation matrix? Are there any sources of data missing from the evaluation matrix that would help answer the evaluation questions? Does the evaluation matrix include sources of data that don't exist or are not accessible?
- **Analysis**: Does the matrix include methods for analysing data in addition to data collection?

Table 7: Process map for inception report quality assurance

Evaluation manager	■ Checks the draft inception report using the Quality Checklist for Inception Report as a guide, to ensure that it follows the structure; provides the main content; and that there are no major contextual and/or factual errors, focusing only on major issues. ■ In case s/he considers the draft inception report does not meet the standards, the evaluation manager provides feedback to the team leader who will submit a revised version.
REU (OEV CapQual unit for HQ DEs)	■ Reviews the draft inception report using the Quality Checklist for Inception Report and confirms that it meets the minimum requirements for submission to DEQS.
Evaluation manager	■ Shares the draft inception report and service request form with the DEQS (copying the Helpdesk, wfp.decentralizedevaluation@wfp.org, the regional evaluation unit (or the Office of Evaluation's CapQual unit for headquarters-led decentralized evaluation) and the chair of the evaluation committee, submitting as well the ToR, the DEQS feedback form for the ToR and how each comment has been addressed and, for those that have not been addressed, the reasoning. An optional call can also be requested if the overall rating is 90 percent or above. ■ As the draft inception report is being reviewed by the DEQS, the evaluation manager and regional evaluation unit (or CapQual unit for headquarters-led decentralized evaluation) conduct a more thorough review against the Quality Checklist for Inception Report and formulate systematic and constructive comments to be consolidated in a comments matrix.
DEQS	■ Acknowledges receipt and confirms return date for feedback within six working days. ■ DEQS project manager may decide to conduct a peer review to ensure that the assessment and feedback by the "reviewer" is appropriate. This will take a few additional days for submitting the final DEQS feedback to the evaluation manager. ■ The regional evaluation unit and the evaluation manager should fill out the survey on the DEQS service at the end of the call.
Evaluation manager	■ For an effective DEQS, the evaluation manager reviews the feedback received from the DEQS before sharing it with the evaluation team leader. In case of issues with the feedback provided by the DEQS, the evaluation manager seeks advice from the regional evaluation unit. ■ Requests a follow-up call with the DEQS reviewer and the evaluation team leader using the DEQS call request form if the overall rating is below 90 percent to clarify any questions arising from the feedback received and to understand the reviewer's overall thinking on the quality of the inception report. The regional evaluation unit (or CapQual unit for headquarters-led decentralized evaluations) should also take part in this call. The call is optional if the inception report is rated above 90 percent.
Evaluation team leader	■ Records in the DEQS feedback form and comments matrix from the evaluation manager and regional evaluation unit how each comment has been addressed and, for those that have not been addressed, the reasoning. The completed comments help ensure a transparent and credible process. ■ Revises the inception report according to feedback from the evaluation manager, regional evaluation unit and the DEQS; submits a second draft inception report to the evaluation manager along with the completed DEQS feedback form with the comments matrix.
Evaluation manager	■ Checks that comments have been adequately addressed and that the inception report is of good quality prior to sharing it with the evaluation reference group for review and comments. ■ Shares draft with evaluation reference group,[21] outlining the previous steps taken to enhance the quality of the draft inception report and providing suggestions about what they should be focusing on (see Box 16).

When the evaluation team is contracted through a consultancy firm, the firm checks the quality of the draft inception report before submitting it to the evaluation manager. Upon receipt of the draft, the evaluation manager follows the steps outlined in Table 7.

In case a baseline is included as part of the contract for the endline evaluation (as opposed to being part of the country office's normal monitoring work), the baseline report should also be submitted to the DEQS. To ease the inception report review process, some practical hints on areas to focus on are provided in Box 15, so that comments focus on the accuracy of information presented.

3.6. Finalize and approve the inception report

What could go wrong?

Evaluation team does not want to address comments after several drafts have been submitted, yet the draft is still not good enough.

Approval of inception report takes too long, further delaying the process.

Time between draft inception report and finalization is too long that context has changed, and the field work needs to be revisited (for example, schools have closed, rainy season started etc).

Good practices

Escalate issues to the long-term agreement firm management and regional evaluation unit (or CapQual in case of headquarters-led decentralized evaluations).

To avoid delays at the approval stage, offer to present the main elements of the inception report to the evaluation committee in a meeting.

Determine which comments are critical (cannot proceed before they are addressed) and focus on those.

On receipt of the revised inception report, the evaluation manager follows the steps outlined in Table 8, consulting the regional evaluation unit (or CapQual unit for headquarters-led decentralized evaluations) as needed.

In terms of data protection, personal, identifiable information of stakeholders attending inception interviews and engaging in the review of the inception report should not be disclosed in the final deliverables published on WFP internal and external websites. Only the affiliation (that is, the organization's units and department) can be reported. Sources of quotations of interviews, focus group discussions and events should not allow for personal identification.

Table 8: Process map for inception report finalization

Evaluation reference group	■ Reviews draft inception report and provides feedback along the questions in Box 16
Evaluation manager	■ Collates all comments in a comments matrix, sorted into groups, according to the section of the report that they relate to. If there are contradictions within the comments, the evaluation manager should discuss how these should be addressed most appropriately with the evaluation team leader. ■ Checks if any of the comments are ambiguous, seeks clarification from the evaluation reference group member. ■ Shares the comments with the evaluation team leader and discusses them with him/her during a meeting or teleconference to deepen mutual understanding of the issues raised.
Evaluation team leader	■ Records in the comments matrix how each comment has been addressed and for those not addressed, the reasoning. The completed comments ensure a transparent and credible process. ■ Revises the inception report and submits the final with the completed matrix of comments to the evaluation manager.
Evaluation manager	■ Reviews the final inception report and the comments matrix with feedback. Any disagreements between WFP, evaluation stakeholders and the evaluation team will need to be addressed and resolved before the inception report is considered finalized and the evaluation can move forward. In case of diverging views with the evaluation team, the evaluation manager consults the regional evaluation unit (or CapQual for headquarters-led decentralized evaluations). ■ Submits the final inception report to the evaluation committee for final review and approval by the evaluation committee chair with offers to present the main elements of the inception report in a meeting to fast track approval.
EC chair	■ Approves the final inception report.
Evaluation manager	■ Shares the approved inception report with the regional evaluation unit, evaluation reference group and all other evaluation stakeholders. ■ Processes payments as per contractual terms and contract management procedure.
REU (OEV CapQual unit for HQ DEs)	■ Shares the approved inception report with the Office of Evaluation for publication on the WFP Office of Evaluation's internal website.

Introduction

	Minimum duration	Recommended duration	Average duration
Duration	0.4 month	0.8 months	1 month
Key outputs	Data sets, de-briefing presentation		

The data collection phase is where the evaluation team collects, synthesizes and starts analysing information and data, from primary and secondary sources as indicated in the inception report. The evaluation team pulls together the evidence to address the evaluation questions.

The details of the data collection phase are determined by the methodology chosen for a given evaluation; therefore, it may differ for each evaluation. The principles provided here are generic but apply to all decentralized evaluations. The data collection phase requires that the inception report is approved. Data collection and analysis tasks are mostly conducted by the evaluation team; however, the country office should support the organization of field visits and end-of-fieldwork debriefing to ensure that they are as effective as possible. Typically, the data collection phase lasts three weeks but it may increase up to one month.

Table 9: Summary of responsibilities – data collection phase

Steps and responsibilities	CD/ DCD	CO Prog. team	CO M&E team	CO Eval Manager	Other CO units	Eval Team	RB	OEV	Other HQ units	External stakeholders
4.1 Prepare evaluation fieldwork		S	S	L (co)	S	L (co)				
4.2 Conduct fieldwork and preliminary analysis	P	S	S	P	S	L	P			P
4.3 Present end-of-fieldwork debriefing(s)	P	P	P	P	P	L	S			P

Key: A: Approve; L: Lead – takes overall responsibility to undertake the step; P: Participate actively in step; S: Support the process with specific tasks, if required

Box 16: Principles of independence, impartiality, credibility and utility

The principles of independence, impartiality, credibility and utility, as well as ethics, have a number of specific implications at data collection phase as follows:

- **Independence**: Ensuring independence and impartiality means that WFP and/or its partner organization staff does not participate in the evaluation team's data collection activities with external stakeholders. The evaluation team leader has the final authority on who should attend data collection meetings. The evaluation manager and the team leader are responsible for anticipating different interests and counteracting attempts to avoid focus on particular issues or particular subgroups, or to influence the evaluation in any way. The evaluation manager ensures that evaluators have full access to available information and data as per the WFP directive on information disclosure. Evaluators have full freedom to conduct their evaluative work without interference or fear for their career.
- **Credibility**: The evaluation manager ensures that the evaluation is implemented as per design. If challenges arise during the field mission, adjustments are made ensuring that they do not undermine impartiality.
- **Utility**: The evaluation team organizes a debrief at the end of the evaluation mission with the support of the evaluation manager and participation of the country office/regional bureau/ headquarters as appropriate.
- **Ethics**: The evaluators behave ethically in all interactions with stakeholders and beneficiaries. They ensure informed consent, protecting the privacy, confidentiality and anonymity of participants, and that cultural sensitivity is maintained. The evaluators respect the autonomy of participants, ensuring their fair recruitment and that evaluation results do not harm them or their communities.

4.1. Prepare for evaluation fieldwork

What could go wrong?

Some elements of fieldwork are underbudgeted/some data collection costs unforeseen till late.

Enumerators' capacity to collect quality data and to implement the procedures is limited.

Poor accessibility of sites: security, poor access roads, weather patterns that affect movement and household activities, network coverage, political instability.

Opaque data quality control processes by evaluators (where they are using enumerators).

Good practices

Insist explicit data quality control measures be put in place (as should have been outlined in the final inception report).

Escalate issues to the firm as early as possible and seek resolution.

Have a contingency plan (budget, sampled sites, targets etc...).

Follow up closely on enumerator trainings, ensure sufficient days are planned for enumerators trainings.

Rely on country office and WFP field/sub-offices for logistical support.

The team leader briefs all team members to ensure they have understood the requirements of the evaluation and the operational plan in the inception report. The team briefing(s) should also serve to clearly agree on the reporting requirements from each team member. The evaluation manager coordinates with the evaluation team to:

- finalize the field site visit agenda in line with the requirements set out in the inception report;
- organize the meetings (including entry briefing and debriefing meetings) and site visits;
- identify the relevant stakeholders to meet;
- prepare the field site visits; and
- provide administrative/logistical support to the evaluation team members including liaising with units/authorities for ticketing, payments, transport, visa, authorizations as relevant.

Communication and coordination mechanisms between the evaluation team and evaluation manager should be established for the duration of the data collection phase and made clear to all at the outset. If the evaluation team is travelling from abroad, this first step is undertaken prior to the arrival of the evaluation team in the country.

Security considerations will vary depending on the country context and the nature of the contracting arrangements with WFP (that is, whether the team is recruited through a long-term agreement or as individual consultants).

As accessibility may be influenced by different factors (security, poor access roads, weather, network coverage, political instability), a contingency plan should be developed to identify alternative sites. The evaluation team can determine a larger sample and request inputs from the country office team to identify those areas where accessibility factors will impede any data collection.

Box 17: Security arrangements for evaluation teams

If the team members are directly recruited by the country office, they are covered by the United Nations Department of Safety and Security (UNDSS) system for United Nations personnel, which covers WFP staff and consultants contracted directly by WFP. Independent consultants must obtain UNDSS security clearance and complete the United Nations system's Basic and Advance Security in the Field courses in advance, print out their certificates and keep them with them.

If the team is contracted through a firm, the contracted firm will be responsible for ensuring the security of the evaluation team, and adequate arrangements for evacuation for medical or security reasons. The WFP commissioning office should, however, take measures to support the security of the evaluation team while in-country. The evaluation manager will ensure that the WFP country office registers the team members with the security officer on their arrival in the country and arranges a security briefing for them to gain an understanding of the security situation on the ground. The evaluation team must observe applicable UNDSS rules including completing online security training (BSAFE & SSAFE) and attending in-country briefings.

4.2. Conduct fieldwork and preliminary analysis

What could go wrong?

Evaluation team does not manage to collect all the information required and/or does not implement the methodology as was designed in approved inception report.

Non-response rate is high due to various reasons.

Data loss (especially where technology such as laptops or tablets are used).

Reported cultural insensitivities by evaluators thus affecting stakeholder responsiveness.

Good practices

Implement the contingency plan (budget, sampled sites, targets...) if required.

Review preparatory work done and ensure evaluation team is accompanied by implementing partner or WFP staff who can make necessary introductions/be gate openers if access is the issue.

Ensure protocols including an off-line plan B to mitigate technology failures.

Keep regular contact with the evaluation team and be available to resolve any issues and/or escalate to the firm.

It is recommended that the evaluation committee chair participates in the entry meeting (or part of it). This meeting with key WFP staff who have a key role in facilitating the evaluation should include:

- a recap on mutual expectations between the country office and the evaluation team;
- the planned agenda for meetings and field site visits (who are the stakeholders, their interests, significance and role in making and/or implementing the subject under evaluation, etc.);
- access to data;
- dates of and expectations for the end of fieldwork debriefing(s). The evaluation team should be given access to the End of Evaluation Debriefing Template;
- expected standards for the evaluation team when collecting data including the UNEG Pledge of Ethical Conduct and Confidentiality Agreement. This includes sensitivity to beliefs and customs, acting with honesty, integrity and respect for stakeholders, and agreement to protect the anonymity and confidentiality of informants; and
- recap on DEQAS quality standards including, but not limited to, GEWE and disability inclusion requirements and their implications at the data collection phase, notably on data collection from both men and women participants, and relevant sub-groups identified as particularly marginalized or excluded where feasible.

The evaluation team starts the data collection work as per the agenda set with the evaluation manager, interacting with WFP and other stakeholders through interviews, focus group discussions, surveys and/or other participatory approaches depending on the evaluation design. As the evaluation team starts data collection, more questions may emerge related to monitoring processes etc. and the evaluation team is likely to request additional secondary data. The M&E team should be available to provide these in a timely manner when available. The list of beneficiaries to be interviewed should be anonymized before being used by the evaluation team.

The evaluation team travels to the selected field sites to carry out further stakeholder interviews, engage with beneficiaries and observe WFP interventions. The duration of these visits may vary depending on the number of activities, partners and ease of access. The evaluation methodology/design is the key determinants of field visit duration. For example, a representative survey would require significant time to conduct and analyse. A typical mission might include one week in the capital city and two weeks visiting relevant area/sub-offices and field sites. The evaluation team may decide to split into separate teams for some activities/regions.

Data protection and confidentiality. The evaluation team must ensure confidentiality, and the protection of data when, processing, storing and/or transferring personal data throughout data collection activities and seek the consent of respondents. More specifically, when seeking personal data through surveys, focus group discussions and interviews, the evaluators should disclose the purpose of the evaluation and with whom the data may be shared. They should also provide contact details of the person/entity to refer to for any concern on the use of his/her personal data. The evaluation team should inform data subjects of their rights including access, rectification, deletion and objection to use. In the course of the evaluation, personal data on beneficiaries should be encrypted and stored to restrict their accessibility. The evaluation team should establish in advance the data retention plan, that is, how long the personal data collected should be available for the use of the evaluation. Should evaluators uncover allegations of wrongdoing or misconduct in the implementation of a programme either by a WFP staff member or a partner, the evaluation team leader should report those allegations to WFP Office of the Inspection and Investigation through the WFP hotline[22].

4.3. Present end-of-fieldwork debriefing(s)

What could go wrong?

Unrealistic expectations from internal/external stakeholders for evaluation team to provide detailed analysis of findings at this stage and/or the evaluation team presents findings prematurely.

Evaluation team members arriving at the debriefing without sufficient agreement between them of what they will present, or voicing diverging views at the debriefings.

Combining internal and external stakeholder debriefings when there are sensitive issues to be presented that the WFP management was not aware of.

Good practices

Explain purpose of the end of fieldwork debriefings to all stakeholders in advance (when inviting them).

Schedule a second debrief later when the evaluation team has had time to analyse the data and have preliminary findings.

Ideally have an initial internal debriefing with WFP stakeholders (except for joint evaluations).

The purpose of the debriefing sessions is to provide an opportunity for the evaluation team to further engage with stakeholders, encourage reflection and validate the preliminary findings. The evaluation team should prepare and present a debriefing, based on data gathered and early analysis conducted. This should cover the team's initial findings/early analysis, any information gaps and next steps. The presentation serves as a reference document to stakeholders but is not reviewed, commented on or revised. The evaluation team should be asked to keep the presentation concise using the End of fieldwork Debriefing Template. Recommendations are not presented at this stage. It is good practice to allow one or two days without any meetings to allow the evaluation team to prepare the debriefing.

Depending on what was agreed during the inception phase, there may be up to three debriefing sessions held. The plan for the debriefings may be revisited after the team has conducted the preliminary analysis.

- **Bilateral meeting with WFP management**: A meeting between the evaluation team and key WFP managers, held before the full internal debriefing to discuss the initial evaluation findings and provide management insights.
- **Internal debriefing**: The evaluation team presents an internal debriefing to key WFP stakeholders (and partners for a joint evaluation), usually held at the main WFP office, but with the option of teleconference participation (for example, for relevant WFP area/sub-office and regional bureau staff).
- **External debriefing**: The evaluation team presents an external debriefing to key external stakeholders (including all evaluation reference group), having absorbed feedback from the internal briefing participants.

The evaluation manager is responsible for providing a venue and facilities; sending invitations well in advance to the evaluation committee, evaluation reference group and other relevant stakeholders; and sharing the debriefing presentation with all stakeholders ahead of the meeting, including the regional evaluation unit (and CapQual for headquarters decentralized evaluations).

Box 18: Recommended content of debriefing presentation

- Purpose (evaluation objectives and uses)
- Context (country/area and operational)
- Methodology overview
- Any changes from the inception report that were made or need to be made as mitigating strategies
- List of field sites visited and rationale
- Stakeholders interviewed and rationale
- Preliminary findings by evaluation question
- Emerging themes for further analysis
- Remaining knowledge/data collection gaps
- Next steps
- Evaluation reference group comments and questions

Introduction

	Minimum duration	Recommended duration	Average duration
Duration	1 month	2.8 months	6 months
Key outputs	Final evaluation report; completed matrix with stakeholder comments		

The reporting phase brings together the evaluation team's findings in a concise, analytical evaluation report, which is the main output of the reporting phase. The reporting phase starts once the evaluation team has completed data collection activities and conducted early analysis. The reporting phase should take approximately 11 weeks if the evaluation team and evaluation manager follow the guidance presented in the guide and there are no unforeseen events. Experience to date shows that it can take longer due to different factors such as the scope and purpose of the evaluation, timely feedback from stakeholders and the quality of the draft evaluation report. There are opportunities to reduce this by three weeks if a good first draft evaluation report is received and shared simultaneously with DEQS and the evaluation reference group.

While reporting is mostly undertaken by the evaluation team, other evaluation stakeholders must feed into the process as shown in Table 10 to ensure quality and utility of the evaluation report.

Table 10: Summary of responsibilities – reporting phase

Steps and responsibilities	CD/ DCD	CO Prog. team	CO M&E team	CO Eval Manager	Other CO units	Eval Team	RB	OEV	Other HQ units	External stakeholders
5.1 Prepare draft ER						L				
5.2 Quality assure draft ER (using DEQS feedback and comments matrix)				L		P	P	P		
5.3 Circulate the draft ER to stakeholders for comments	P	P	P	L	P		P		S	P
5.4 Finalize and approve evaluation report	A	P	P	L		L	S			

Key: A: Approve; L: Lead – takes overall responsibility to undertake the step; P: Participate actively in step; S: Support the process with specific tasks, if required

Box 19: Principles of independence, impartiality, credibility and utility

The principles of independence, impartiality, credibility and utility, as well as ethics, have a number of specific implications at the reporting phase as follows:

- **Independence**: The evaluation team must be free from pressure to alter conclusions and recommendations in any way that is not supported by the evaluation's findings.
- **Impartiality**: The evaluation team analyses data and presents findings transparently and reflects where different stakeholders held different views while ensuring confidentiality. The evaluation team must provide explicit rationale when it does not incorporate stakeholder feedback.
- **Credibility**: Requires that findings are based on triangulated evidence: clear connections are made between findings, conclusions and recommendations. Findings and conclusions are fair and acknowledge the existence of differing views. The evaluation report explains the methodology and any limitations; and presents evidence, findings, conclusions and recommendations in a complete and balanced way.
- **Utility**: With the support of the evaluation manager, the evaluation team ensures an efficient reporting process as per the evaluation timeline to avoid late completion of the evaluation. A stakeholder workshop is organized to present initial findings, conclusions and recommendations to internal stakeholders.
- **Ethics**: All evaluations should reflect whether the ethical risks identified in the inception report (Section 3.5) materialized and, if they did, how and what safeguards were employed to mitigate them. Should evaluators uncover allegations of wrongdoing and misconduct in the implementation of a programme, either by a WFP staff member or partner, the evaluation team leader should report the allegations to the WFP Office of Inspection and Investigation through the WFP hotline. At the same time, the evaluation team leader should inform the evaluation manager, country office management, the regional evaluation unit and the Director of Evaluation that there are allegations of wrongdoing and misconduct, without breaking confidentiality.

5.1. Prepare the draft evaluation report

What could go wrong?

Team realizes that they have significant gaps in data and cannot answer some questions.

Delays in drafting and submitting the evaluation report.

Evaluation team does not use the WFP template for evaluation reports.

Evaluation team submits a poor quality/ incomplete draft evaluation report.

Good practices

Keep regular contact with the evaluation team, discuss the nature of data gaps and see if they can be filled with secondary data. If not the evaluation team should report limitations and which questions cannot be answered.

Return incomplete evaluation report to the team and require that they use appropriate template and guidance.

Keep stakeholders informed of the progress and any delays.

Remind the firm of WFP expectations that they quality assure the draft before submission to WFP.

The evaluation team leader prepares the draft evaluation report with the support of the evaluation team. The evaluation report should follow the structure and include the content provided in the Evaluation Report Template and should meet DEQAS quality standards (see Quality Checklist for Evaluation Report and the Technical Note on Quality of Evaluation Recommendations).

Data privacy and protection considerations: the evaluation report should not use personal data, not profiling data subjects, data anonymization, security of personal data, accuracy of personal data used for analysis, retention of personal data only for the time that is necessary for the specified or compatible evaluation purposes. Once the purpose has been fulfilled, personal data shall be anonymized, deleted, or erased.

Joint

For **joint evaluations**, the set of recommendations should offer opportunities for programme enhancement for all commissioning agencies. Some recommendations from a joint evaluation may target an individual agency, others refer to all commissioning agencies (possibly with disaggregated sub-recommendations). If the evaluation policy of any of the partners involved does not require evaluations to make recommendations, WFP will still require recommendations. This should be made clear at the start of the process.

Box 20: Process for formulating and increasing ownership of recommendations

Engaging with evaluation users as early as possible in the formulation of recommendations will contribute to more relevant and actionable recommendations, as well as increase the level of ownership of the evaluation results by key users. Therefore, in addition to seeking written feedback on the draft report via email, it is suggested that a meeting/workshop is organized to:

- reflect on key evaluation findings and conclusions;
- discuss the draft recommendations and provide feedback to the evaluation team to enhance their clarity, relevance and feasibility; and
- start reflecting on how the country office will address the recommendations and the management response.

5.2. Quality assure the draft evaluation report

What could go wrong?

Poor quality draft evaluation report even after returning initial draft to the evaluation team for revisions.

Significant delays in submission of draft evaluation report putting pressure on quality assurance timelines especially if there are time-sensitive decisions to be made based on the results.

DEQS feedback not received on time further delaying the process.

DEQS feedback not useful.

DEQS feedback useful but not used to improve the quality of the evaluation report.

Good practices

Allocate in advance sufficient time to review the draft evaluation report (if possible, blocking a full day).

Keep DEQS calendar/tracker updated so the service provider lines up the reviewer.

Discuss with evaluation team whether a presentation of findings and recommendations could be made to key stakeholders while waiting for the final report.

When the evaluation team is contracted through a consultancy firm, the firm checks the quality of the draft evaluation report before submitting it to the evaluation manager. Upon receipt of the draft evaluation report, the evaluation manager follows the steps outlined in Table 11.

Table 11: Process map for evaluation report quality assurance

Evaluation manager	■ Reviews the draft evaluation report using the Quality Checklist for Evaluation Report as a guide, to ensure that it follows the structure; provides the main content; and that there are no major contextual and/or factual errors. ■ In case s/he considers the draft evaluation report does not meet the standards outlined above, the evaluation manager provides feedback to the team leader so that the report adheres to the quality standards.
REU/ CapQual	■ Reviews the draft evaluation report using the Quality Checklist for Evaluation Report and confirms that it meets the minimum requirements for submission to the DEQS. ■ Takes part in the call with DEQS reviewer.
Evaluation manager	■ Shares the draft evaluation report and service request form with the DEQS (copying the Helpdesk, wfp.decentralizedevaluation@wfp.org, the regional evaluation unit and the chair of the evaluation committee), together with the inception report, DEQS feedback form for the inception report that should have been completed by the evaluation team before submission of final inception report. ■ As the draft evaluation report is being reviewed by the DEQS, the evaluation manager and regional evaluation unit conduct a more thorough review of the draft report against the Quality Checklist for Evaluation Report and formulate systematic and constructive comments to be consolidated in a comments matrix.
DEQS	■ Acknowledges receipt and confirms the return date for the feedback within six working days. ■ DEQS project manager may decide to conduct a peer review to ensure that assessment and feedback by the "reviewer" is appropriate. This will require additional days for submitting feedback. ■ The regional evaluation unit and the evaluation manager fill in the survey on the DEQS service at the end of the call.
Evaluation manager	■ Reviews the feedback received from the DEQS before sharing it with the evaluation team leader. In case of issues with the feedback provided by the DEQS, the evaluation manager seeks advice from the regional evaluation unit. ■ Requests a follow-up call using the DEQS call request form, with the DEQS reviewer and the evaluation team leader to clarify any questions arising from the feedback received and understand the reviewer's overall thinking on the quality of the evaluation report. The regional evaluation unit (or CapQual unit for headquarters decentralized evaluations) and evaluation team leader should also take part in this call. This call is optional for evaluation reports rated above 90 percent.
Evaluation team leader	■ Records in the DEQS feedback form and comments matrix from the evaluation manager how each comment has been addressed and, for those that have not been addressed, the reasoning. The completed comments help ensure a transparent and credible process. ■ Revises the evaluation report according to feedback from the evaluation manager, the regional evaluation unit and the DEQS; submits a second draft to the evaluation manager along with the completed DEQS feedback form with the comments matrix from the evaluation manager.
Evaluation manager	■ Ensures that comments have been adequately addressed and that the evaluation report is of good quality prior to sharing with the evaluation reference group for review along with the comments matrix.

STEP 1
STEP 2
STEP 3
STEP 4

5.3. Circulate the draft evaluation report to stakeholders (for comments)

What could go wrong?

Very little or no comments/feedback from stakeholders.

Stakeholders' inputs contradict each other.

Stakeholders seeking to unduly influence findings/conclusions especially if they are negative.

Evaluation reference group/stakeholders take too long to provide feedback or provide after process has moved on.

Good practices

Establish in advance a circulation list (in addition to evaluation reference group members) in collaboration with the evaluation team and regional bureau focal points and discuss the modalities and timing to elicit stakeholder comments.

Allow sufficient time – two weeks is the norm – for stakeholders to read and provide comments in writing on the draft report. Sending a reminder a few days before the deadline can be useful.

Provide clear guidance on how stakeholders should share their comments and let people know that their comments will be consolidated in a matrix of comments, responded to and circulated back to all reviewers to encourage considered feedback.

Offer to present the main findings, conclusions and recommendations in an evaluation reference group meeting to elicit feedback.

The evaluation reference group is normally given two weeks to review and provide comments on the draft evaluation report. Exceptionally, this may need to be shortened to five–eight working days in agreement with the reviewers to ensure that the report is available to decision makers when required. When communicating with the evaluation reference group, the evaluation manager should describe the previous steps taken to enhance the quality of the evaluation report.

Joint

The risk that the quality assurance process becomes overly cumbersome is high in the case of **joint evaluations**. Mitigation strategies need to be pursued, such as the rationalization of the number of evaluation reference group members; the organization of multi-stakeholder review workshops; or the establishment of cascade review systems to avoid evaluation managers receiving an excessive number of comments.

To ease the report review process, the evaluation manager could provide guidance to reviewers so that their comments focus on the accuracy of information presented. See below for some examples:

- Are the data accurate? If not, please provide additional materials for consideration.
- Are there significant gaps in the data presented?
- Are the data presented in a clear, unbiased and objective manner?
- Are the findings consistent with the data presented?
- Do the conclusions logically follow the findings?
- Are the recommendations feasible and do they follow logically from the findings and conclusions?

The evaluation manager should consolidate all comments into a comments matrix using the template provided, sorted into groups, according to the section of the report that they relate to. If there are contradictions between several comments, the evaluation manager should discuss how these should be addressed with the evaluation team. If any of the comments appear ambiguous, the evaluation manager should seek clarification from the reviewer first before sharing them with the evaluation team leader.

Time invested at this stage helps improve the quality of the draft evaluation report, making for a more efficient reporting process. A dialogue between the evaluation manager and the evaluation team is required to:

- enhance mutual understanding of the feedback received from evaluation reference group members;
- provide clarification on specific areas when requested; and
- discuss the main areas of findings, conclusions and recommendations as they are developed.

The evaluation team must consider comments received but has the discretion to accept or reject stakeholder comments and any requests for changes. Where there is pressure to alter conclusions or recommendations, while still consistent with findings, such requests should be judged on their merits. Should any differences arise, which cannot be resolved through discussions between the evaluation manager and team leader, the evaluation reference group should be asked to intervene, and contact the Office of Evaluation decentralized evaluation help desk as appropriate. In case of diverging views within the evaluation team, the evaluation manager should consult the regional evaluation unit and the Office of Evaluation for advice.

5.4. Finalize and approve the evaluation report

What could go wrong?

Some stakeholders send comments long after the commenting stage has passed.

The final report is still not of good quality even after stakeholder comments, including: having recommendations that are not in line with findings and conclusions; conclusions that are a copy and paste of findings; and that the evaluation report lacks a lessons learned section even although drawing lessons was an objective.

The evaluation committee chair doesn't approve the final report on time, further delaying and risking reduced utility (or the evaluation committee does not want to approve if the evaluation team does not make certain changes).

Good practices

When sharing products for comments, always do so on a 'no-objections basis', where no comments received after deadline will be considered to mean no feedback unless a deadline extension is sought.

Review the draft evaluation report to ensure consistency across the analysis.

Return both the report and comments matrix as many times as required until the evaluation team either addresses all comments OR provides a proper justification for not addressing the comment.

Engage with the evaluation committee chair and escalate when needed.

Present the findings, conclusions, and recommendations in an evaluation committee meeting to facilitate approval.

A completed comments matrix helps ensure a transparent and credible process. It is important that the evaluation team leader acts on stakeholders' comments to correct factual errors, avoid misrepresentation and provide additional information as the evaluation report is finalized. The evaluation team should exercise judgement as to which comments merit follow-up. For any comments not addressed, the evaluation team should provide adequate rationale. This supports transparency and credibility.

If the evaluation manager finds that some of the comments have not been duly considered, the evaluation team should be asked to address the comments and resubmit the final report and comments matrix. Any disagreement between WFP, evaluation stakeholders and the evaluation team should be addressed and resolved before the evaluation report is considered finalized. The evaluation manager can consult the regional evaluation unit or the Office of Evaluation for advice.

In terms of data protection, personal identifiable information of stakeholders attending inception interviews and engaging in the review of the evaluation report should not be disclosed in the final deliverables published on WFP internal and external websites. Only affiliation (that is, organization units/department) can be reported in the documents. Sources of quotations of interviews, focus group discussions and events should not allow for personal identification.

To ensure standardization and compliance with the corporate communications policy, the formatting guidelines should be followed when finalizing the evaluation report.

Joint

For **joint evaluations**, coordination of feedback to draft evaluation deliverables risks becoming very complex. For controversial issues, bilateral preparatory meetings with key decision makers among the participating agencies should take place in advance of the formal consultation. Co-evaluation managers (or the evaluation management group) and country office management have a role to play in ensuring that the concerns of the commissioning agencies have been taken into consideration while not undermining the impartiality of the evaluation. Ample time to review the final report should be allowed, requesting comments on a 'non-objection' basis (no response = agree) so that the process is not halted for an unnecessarily long period of time. It is important to ensure that written approval of the evaluation report is provided by all members of the evaluation reference group, circulating a PDF file that cannot easily be modified, to ensure ownership of all the commissioning agencies at evaluation dissemination and follow-up stages.

McGovern Dole

Evaluations of McGovern Dole-funded school feeding programmes must be 508-compliant[23] so that they are accessible for persons with disabilities. Evaluation managers of these types of evaluations should be aware of additional reporting requirements (and commenting stages) and plan accordingly in close collaboration with the school-based programmes MEAL team.

Table 11: Process map for evaluation report quality assurance

Evaluation team leader	■ Records in the comments matrix how each comment from the evaluation reference group has been addressed and, for those that have not been addressed, the reasoning. The completed comments ensure a transparent and credible process. ■ Revises the evaluation report accordingly and submits the final and the completed matrix of comments to the evaluation manager.
Evaluation manager	■ Reviews the final evaluation report and the comments matrix with the evaluation team's feedback. ■ Submits the final evaluation report to the evaluation committee for final review and approval by the evaluation committee chair.
EC chair	■ Approves the final evaluation report.
Evaluation manager	■ Shares the approved evaluation report with the regional evaluation unit and ensures the final report is proactively shared with the evaluation reference group and all evaluation stakeholders. ■ Processes payments as per contractual terms and contract management procedure.
REU/ CapQual	■ Shares the final evaluation report with the Office of Evaluation for publication on the WFP internal and external evaluation websites.

Introduction

	Minimum duration	Recommended duration	Average duration
Duration	0.7 month	2 month	2.2 months
Key outputs	Management response to recommendations; learning products		

This section provides an overview of the final steps in the evaluation process. The dissemination and follow-up phase is key to making sure that the findings, conclusions and recommendations of the evaluation are made available to the intended audience so that they can be used to help improve WFP interventions, strategies and policies. Timeframes for this phase will vary, but will be around eight weeks.

The dissemination and follow-up phase aims to maximize the use of the evaluation findings by:

- developing a management response (MR) to the evaluation recommendations;
- disseminating the evaluation report and management response to a wide audience; this may include the development of learning products tailored to specific audiences (as per the communication and knowledge management plan);
- submitting the evaluation report for post hoc evaluation quality assessment (PHQA); and
- identifying opportunities/entry points for promoting the uptake of evaluation recommendations.

Table 13: Summary of responsibilities – dissemination, follow-up and use phase

Steps and responsibilities	CD/ DCD	CO Prog. team	CO M&E team	CO Eval Manager	Other CO units	RB	OEV	Other HQ units	External stakeholders
6.1 Prepare management response & upload in R2 system	L	P		P	P	L	S		
6.2 Publish ER and WFP MR				P		P	L		
6.3 Submit ER for PHQA						L	P		
6.4 Hold end-of-evaluation lessons learned debriefing	P			P		L	P		
6.5 Disseminate and use evaluation results	L	L	L	L	L	L	L	L	L
6.6 Track implementation of follow-up actions to the evaluation recommendations in R2 system	P	P	L	L (CO R2 focal point)	P	P		S	

Key: A: Approve; L: Lead – takes overall responsibility to undertake the step; P: Participate actively in step; S: Support the process with specific tasks, if required

Box 21: Principles of independence and utility

The principles of independence, impartiality, credibility and utility, as well as ethics, have a number of specific implications at the dissemination and follow up-phase, as follows:

- **Independence**: All final evaluation reports, management response and post hoc quality assessment results are published on WFP websites and disseminated through various channels.
- **Utility**: A management response is prepared for all evaluations detailing how the evaluation recommendations will be addressed. The implementation of follow-up actions is monitored. Opportunities for wider organizational learning are pursued including incorporating key discussions on evaluation results in key workshops and programme, strategy or policy-making processes.

6.1. Prepare management response and upload in R2 system

What could go wrong?

Process takes too long especially if it is started at the time the evaluation report is approved.

Stakeholders misunderstand the meaning of "disagreeing with recommendations" and disagree with those they have already acted on and require no further action or those for which they disagree with the recommended timelines.

Actions are too general or too vague

Good practices

Start engaging stakeholders in thinking about the responses to the recommendations during the finalization of the evaluation report

Include a stakeholder workshop with the evaluation team and stakeholders before the evaluation report is approved.

Provide support to conduct a management response workshop to elucidate the actions that will be taken.

Review the draft management response and explain to stakeholders the conditions under which they should disagree with recommendations as per guidelines.

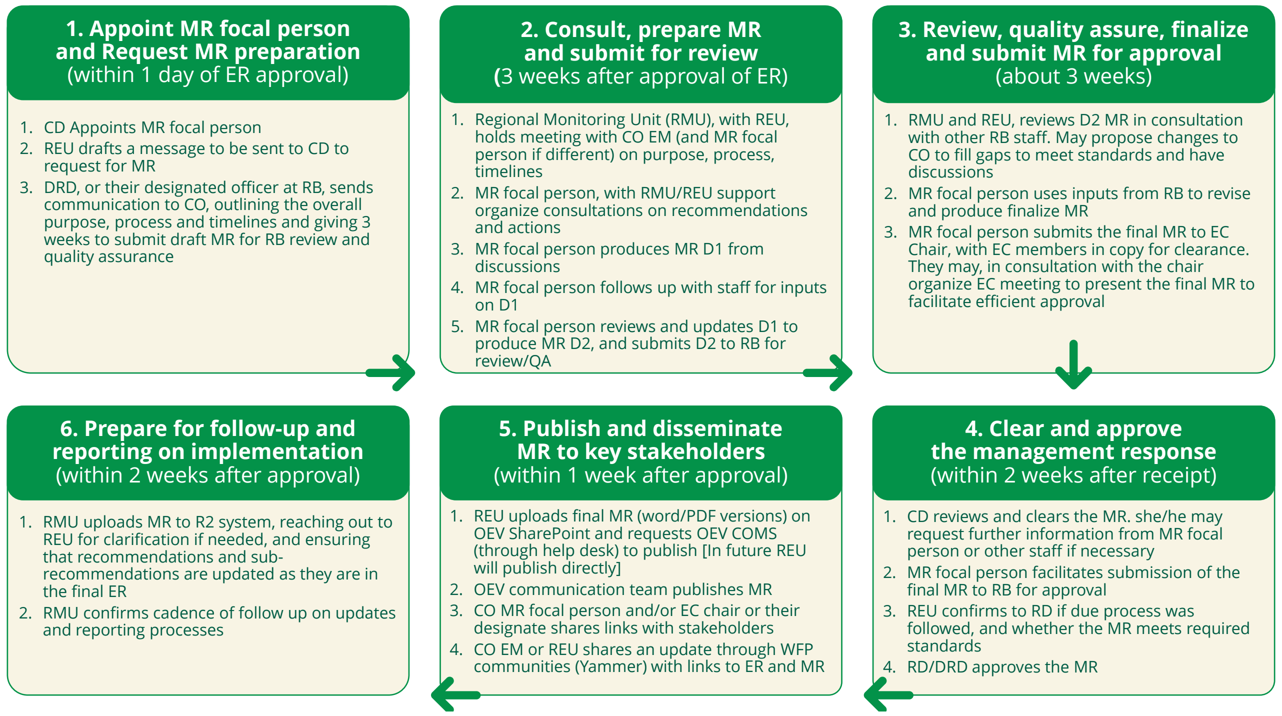
The management response specifies the actions that the commissioning office has agreed to take to address each evaluation recommendation, along with the corresponding implementation deadlines. The management response should be prepared within eight weeks of the approval of the final evaluation report. To enhance accountability and learning, the evaluation recommendations and corresponding management response should be entered in the WFP Risk and Recommendation Tracking Tool (R2) system (the WFP corporate system for management responses), and the responsible office should track and update the implementation status of recommendation accordingly.

The Technical Note on Management Response to Decentralized Evaluation Recommendations provides guidance on the timelines, process and roles and responsibilities for the preparation of a management response for decentralized evaluations. These are summarized in Figure 17.

For **joint evaluations**, a joint management response to evaluation recommendations should be prepared, covering all the recommendations with specific agencies as being responsible for implementation as appropriate. Agreed follow-up actions should be agency-specific to ensure adequate commitment for their implementation. Some partner evaluation guidance may not require management response. However, WFP will always require a management response for evaluations it is part of. In such cases, WFP and the partners that require a management response will jointly prepare it.

For **joint evaluations**, only actions for which WFP is fully responsible are entered into R2. However, to ensure optimal use of joint evaluations, it is important to make sure that the commissioning agencies define the responsibilities and frequency for tracking the follow-up actions related to all recommendations. This could for example be done during regular programme coordination meetings relevant to the subject of evaluation.

Figure 17: Process map for the preparation of management response to decentralized evaluation recommendations



6.2. Publish evaluation report and WFP management response

The evaluation manager is responsible for sharing the final evaluation report, management response and other learning products with the regional evaluation unit, who then shares it with the Office of Evaluation's communication team for publication on internal and external WFP websites; and for posting the report on the WFP Evaluation Community.

When sharing the evaluation report for publication, the report should be accompanied by an evaluation "blurb" to introduce and describe the evaluation product on the website. This helps to generate interest and facilitate the visibility and announcement of the publication.

Box 22: Evaluation "blurb" template

This decentralized evaluation was commissioned by the country office and covers the 'topic' and "evaluation period" (year-year). It was carried out in 'year'. The evaluation was commissioned to (for example, fill an evidence gap, inform the scaling up of a pilot, inform the design of the country strategic plan, etc.) and was intended for both accountability and learning purposes. It focused on assessing the relevance, effectiveness, efficiency, impact, sustainability, coherence and connectedness of the X (edit as required selecting only the relevant evaluation criteria). Overarching evaluation questions focused on xxxx (insert main top-level overarching questions). The evaluation covered the following activities (or themes depending on the evaluation type): a) school feeding; b) nutrition; c) food assistance for assets; d) food and nutrition security analysis & diagnostic; and e) capacity strengthening. Key evaluation findings included (add maximum five bullet points). Key recommendations from the evaluation included (add maximum five bullet points).

STEP 1
STEP 2
STEP 3
STEP 4
STEP 5
STEP 6

6.3. Submit evaluation report for post hoc quality assessment

In line with the WFP Evaluation Policy's provisions of independence and impartiality, the Office of Evaluation ensures independent post hoc quality assessment (PHQA) of all evaluations, contributing to enhancing the quality and credibility of WFP evaluations and strengthening organizational accountability and transparency.

Independent evaluation experts assess the quality of completed evaluation reports against DEQAS quality standards, including the requirements for evaluation set by the United Nations System-Wide Action Plan on Gender Equality and the Empowerment of Women (UN-SWAP) and United Nations Disability Inclusion Strategy (UN-DIS). As such, the PHQA should not be regarded as a quality assurance mechanism, because it comes at the end of the evaluation process and does not contribute to the improvement in quality of draft evaluation reports. However, the PHQA contributes to the quality of future evaluation reports as it helps identify good practices to share, as well as areas for improvement to address through the DEQAS and capacity development initiatives.

Box 23: Post hoc quality assessment criteria

PHQA is designed to: assess the extent to which the WFP evaluation reports are based on appropriate evaluation methods; present a sound analysis and credible findings; and provide valid conclusions and useful recommendations. Quality is assessed against eight criteria, with gender, equity and inclusion elements mainstreamed throughout:

1. Report summary
2. Context and overview of the evaluation subject
3. Rationale, objectives and scope
4. Methodology
5. Findings
6. Conclusions and lessons learned
7. Recommendations
8. Accessibility and clarity

Each criterion is associated with a score reflecting its relative weight in determining the quality of a report. The PHQA tool also provides a score for the report's integration of gender, as per the standards outlined in the UN-SWAP Evaluation Performance Indicator (UN-SWAP EPI) and the United Nation Disability Inclusion Strategy Evaluation Performance Indicator (UN-DIS EPI).

Table 14: Process map for the post hoc quality assessment

REU (OEV CapQual unit for HQ DEs)	■ Uploads the ToR, inception report, final evaluation report and annexes in Word in the PHQA Teams folder ■ Enters the email addresses of recipients of the PHQA summary
OEV	■ Shares the draft PHQA results with the regional evaluation unit for feedback within six weeks of the evaluation report approval. ■ Shares final PHQA results with commissioning office, publishes them on WFP internal and external websites, and consolidates all PHQA results in the annual evaluation report. ■ On receipt of the post hoc quality assessment results, shares final evaluation reports with UNEG and ALNAP (unless the report was rated as unsatisfactory). ■ Shares summary results of the PHQA with the evaluation team that conducted the evaluation

6.4. Hold end-of-evaluation lessons learned debriefings

End-of-evaluation lessons learned debriefing sessions are an opportunity to jointly reflect and discuss the completed evaluation, along with identifying and documenting key lessons on the process. These sessions are also intended to inform the commissioning office of future evaluations, adjustments to the DEQAS guidance and WFP support mechanisms for decentralized evaluations.

The regional evaluation unit should coordinate with the country office evaluation manager (and CapQual with the headquarters evaluation manager) to schedule and facilitate these discussions by sharing the questions with participants ahead of the debriefing. The Office of Evaluation has developed a set of guiding questions to facilitate these specific discussions.

These should take place in two rounds:

- Once the management response is approved: a lesson learned call should be scheduled with the evaluation manager, the chair of the evaluation committee, and, if available, the Office of Evaluation CapQual unit. If the evaluation committee chair is not available, the call can take place with only the evaluation manager (or include the head of evaluated subject) to avoid delays.
- Once the PHQA is finalized: a lessons learned call should be scheduled with the evaluation manager, the team leader, the firm's quality assurer (if possible) and the long-term agreement focal point. The regional evaluation unit is responsible for circulating the PHQA results ahead of the meeting and a set of guiding questions to facilitate these specific discussions. This will help to reflect on the evaluation process and results

The regional evaluation unit is responsible for uploading into the Office of Evaluation's SharePoint the note for record of the calls and sharing it with the participants, for information.

6.5. Disseminate and use evaluation results

It is important that evaluation reports are accessible to a wide audience, as foreseen in the Evaluation Policy, to ensure the credibility of WFP through transparent reporting and the usefulness of evaluations.

The Communication and Knowledge Management Plan sets out who is responsible for each dissemination activity, what material is to be disseminated (for example, the full evaluation report, the executive summary only, an evaluation brief, a video etc.), to whom, how, when and why. Systematically planning for the dissemination of the evaluation results is the best way to ensure evaluations go beyond a mere internal exercise. These dissemination activities should be reviewed and enacted during this phase.

Evaluation results should be used to inform and improve future programme design and implementation, advocacy and strategy and, in particular, evaluation evidence should be used for the design of new CSPs and/or revisions. It is therefore important to actively share lessons among stakeholders, including all donors that contributed to the intervention being evaluated, and with other colleagues in the country office, regional bureau and headquarters to inform similar interventions in other places.

To maximize the use of the evaluation findings, different products should be developed for different groups of users depending on their information needs. The evaluation manager should consider and use audience-appropriate means for sharing the evaluation results so that stakeholders understand and participate in plans to act on recommendations. For example, language, internet accessibility and connectivity issues need to be explored when matching the type of product to the audience. Evaluation dissemination should specifically consider GEWE and disability inclusion dimensions. A GEWE-responsive dissemination strategy might include: (i) dissemination of evaluation findings on GEWE to diverse groups of stakeholders who have an interest in, and are affected by, GEWE issues; and (ii) promoting the fullest possible use of GEWE issues of the evaluations within the United Nations system, NGO partners and the government ministries among stakeholders. Alternative ways to present GEWE-responsive evaluation findings to women and individuals/groups who are marginalized should be considered, including disability inclusive and accessible communications.

Box 24 highlights effective ways to disseminate evaluation evidence and results, and best practices on enhancing the use of evaluation.

Box 24: Effective ways to disseminate evaluation evidence and results

- Send the full evaluation report to the evaluation committee and the evaluation reference group.
- Circulate a stand-alone executive summary or a one-pager evaluation brief to other internal and external stakeholders (along with a link to the full report) as identified in the Stakeholder Analysis and Communication and Knowledge Management Plan.
- Hold workshops in the country – interactive events are more likely to be effective in disseminating evaluation findings.
- Identify and use other emerging opportunities for sharing or using the evaluation findings and recommendations, for example consultation meetings at country office, regional bureau or headquarters levels.
- Identify creative methods to further disseminate the evaluation findings and stimulate discussions including with the affected population, for example, evaluation briefs/pamphlets/ posters/ PowerPoint/ drawings/illustrations/presentations that summarize key evaluation information.
- Making use of audio-visual technology to produce videos and radio clips, short movies, songs, computer animation, community consultation etc. is encouraged. These methods may be particularly useful in reaching stakeholders with low literacy levels. Innovative formats such as graphic novels, cartoons or live storytelling may also be an effective way of disseminating results.

See as examples the following videos:

  - WFP mid-term decentralized evaluation of school feeding in Bangladesh
  - WFP thematic decentralized evaluation on gender in El Salvador
  - Inter-Agency Humanitarian Evaluation of the Drought Response in Ethiopia
  - WFP mid-term decentralized evaluation of the Protracted Relief and Recovery Operation 200708 in Colombia
  - WFP decentralized evaluation of the cash-based transfer modality in school feeding activities in Senegal

- Ask partner organizations to post links to the evaluation report on their websites.
- Use the online Evaluation Community, or other thematic WFP communities, as a channel through which to disseminate evaluation findings.
- Make the evaluation report more accessible to the broader evaluation and humanitarian community by uploading it onto the UNEG and ALNAP evaluation report databases (including the ALNAP evaluation evidence map, Evalmapper).
- Use evaluation evidence in the design of new CSP/Interim CSP and revisions.

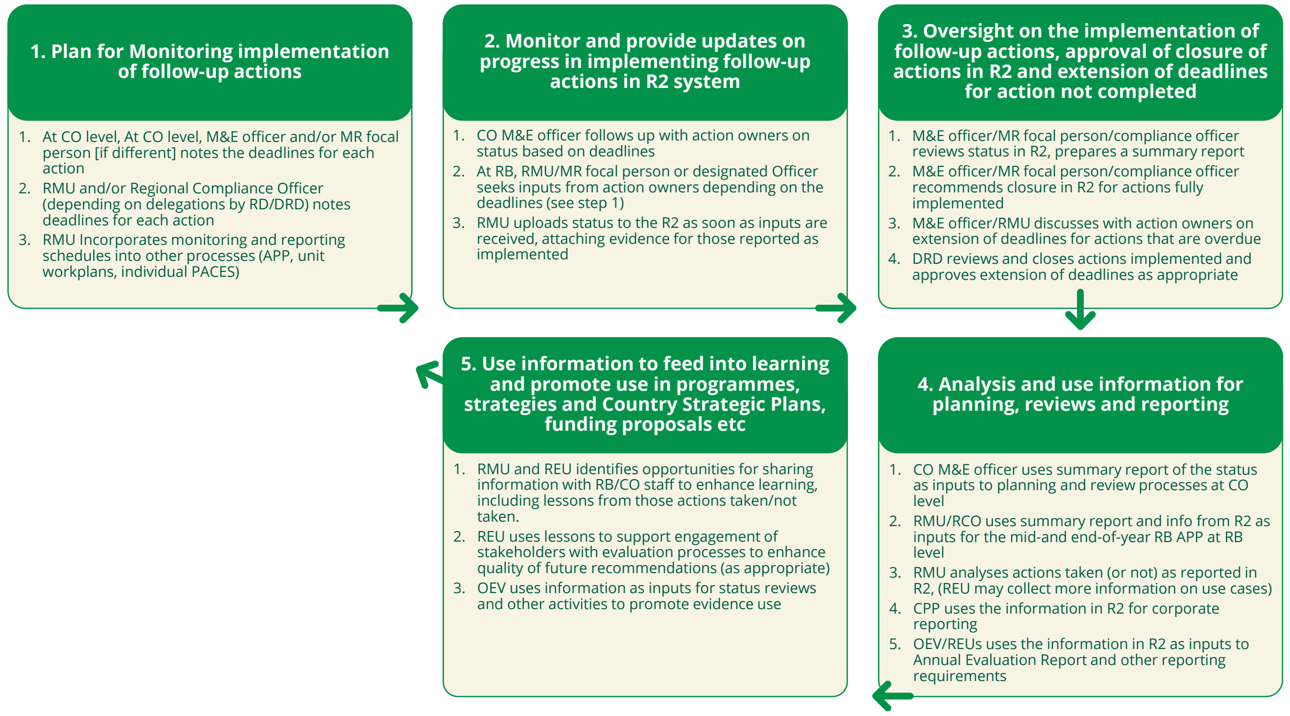
6.6. Track implementation of follow-up actions to the evaluation recommendations in R2 system

WFP management is expected to report on the progress it has made in implementing actions/ recommendations agreed in the management response twice a year. WFP uses the corporate Risk and Recommendation Tracking Tool (R2), managed by the WFP Corporate and Performance Division (CPP), for administering the follow-up process to evaluation recommendations.

The Technical Note on Management Response to Decentralized Evaluation Recommendations provides guidance on the timelines, process and roles and responsibilities for the follow-up on management response actions for decentralized evaluations. These are summarized in Figure 18 below.

Creating spaces for regular conversations around evaluation recommendations, status of implementation and lessons emerging from actions taken is likely to enhance organizational learning.

Figure 18: Process map for the tracking implementation of follow-up actions to the evaluation recommendations



Annex

Key documents and sources of data

QUALITATIVE INFORMATION

Process, guidance and corporate documents	DEQAS Process Guide, templates, quality checklists and technical notes
	Key relevant WFP policies and strategy documents such as evaluation, CSP, gender, protection, nutrition, social protection etc [programme staff should have these to share]
	Key relevant WFP programme guidelines and supporting material (see the programme guidance manual)
	Summary of Evaluative Evidence
	WFP Corporate Results Framework
Country context	National policies and frameworks (national and United Nations):
	Government policies, strategies, assessments and programme documents
	UNSDCF planning documents
	Information on limitations (access and security issues, seasonality etc.)
CSP and operational documents	CSP narrative document (or other documents describing the subject of the evaluation) and subsequent budget revisions (if any)
	NFR from S-PRP meeting and comments submitted through e-PRP on original CSP and subsequent budget revisions
	Annual performance reports and risk registers
	Notes for Record of Coordination meetings (clusters for example, food security, nutrition, logistics; donor groups, etc.)
	Information related to cooperating partners (field-level agreements; memorandums of understanding; lists of partners by activity and location)
	Maps (operational (project sites, distribution sites, warehouse sites), food security, IPC phases)
Human resources	Information on country office structure (number of sub-offices, locations, organigrams, etc.)
VAM, monitoring and evaluation	Key operational documents, including annual country reports (ACRs), post-distribution monitoring reports, internal reports on output and process monitoring, logframes, VAM assessments, studies by the country office and its partners; etc.
	Dispatch and distribution reports from WFP and partners
	Previous evaluations or reviews (thematic or CSP mid-term reviews)
	Extract from R2 system on status of implementation of follow-up actions from past evaluations Programme reports (For example, Pipeline, Projected Needs)

QUANTITATIVE INFORMATION

Type	Quantitative data	Possible analysis
Country context	WFP Dashboard, provides data on: ■ Chronic hunger (number of people) ■ IPC/CH (number of people in phases 3+, 4 and 5) ■ Nutrition (prevalence of wasting/ stunting in children) ■ Displaced people (number of refugees and IDPs) **From relevant national statistics office**: ■ Poverty data ■ Population data ■ Most recent household (HH) census implemented in the country	■ Provide a general overview of the food and nutrition security situation in the country and on migration, refugees and internally displaced people. ■ Provide a general overview on socioeconomic indicators in the country (poverty, inequality, health etc.)
Budget and implementation plan	Budget and implementation plan: ■ Country portfolio budget (original and budget revisions if any) ■ Implementation plan	■ Analysis of the budget by strategic outcomes (SOs)/ activities ■ Analysis of prioritization of resources: comparison between needs-based plan and implementation plan by strategic outcomes/activities
Financial contributions and resources updates	FACTory, CSP Resource Situation report: shows allocated contributions by donor	■ Analysis of key CSP donors
	FACTory, Weekly Publication Reports: customizable pivot table linked to the detailed contributions and forecast dataset[24]	■ Analysis of donor contributions by earmarking level/WBS ■ Analysis of donor contributions by focus area
	IRM Analytics, ACR1 - Annual Country Report - Cumulative Financial Overview: covers from the start of the CSP till the selected date. It provides NBP, Allocated Contributions, Expenditures by activity, SO, and focus area	■ Cost efficiency analysis: Disbursement rates (expenditure vs. mobilized) per cost category (total direct costs, DSC, ISC, overall budget), per year ■ Cost efficiency analysis: Disbursement rates (expenditure vs. mobilized) per SO and activity, per year ■ Cost efficiency analysis: Direct support costs/total direct costs ratio, per year
	IRM Analytics, ACR5 - Annual Country Report - Annual Financial Overview: for each CSP year, provides NBP, Allocated Contributions, Expenditures by activity, SO and focus area	
	IRM analytics: CPB – Plan vs Actuals Report: NBP, Implementation Plan, Actual by activity, transfer modality and by year	■ Cost efficiency analysis: Cost per beneficiary per transfer activity, planned vs. actual, per year
		■ Cost efficiency analysis: Cost per MT of food distributed per year, compared to other humanitarian actors in the same areas, WFP operations in comparable countries and regional averages
		■ Cost efficiency analysis: Cost per USD of cash distributed per year, compared other humanitarian actors in the same areas, to comparable countries and regional averages

Type	Quantitative data	Possible analysis
Human resources	WFP Dashboard, provides data on: ■ Total number of employees ■ Number of employees disaggregated by gender ■ Number of employees by contract duration (long/ short term) ■ Number of international and national employees ■ IRM analytics; CORA staff report, provides data on staffing costs and expenditures.	■ Staffing analysis ■ Staffing costs: planned vs. actual, expenditures by cost category and by activity, and by staff type.
Beneficiaries	COMET, Annual figures: ■ CM-R001b – Annual Country Beneficiaries (CSP): Actual vs planned beneficiaries by gender, residence status, age group	■ Yearly actual vs planned beneficiaries by gender ■ Yearly actual vs planned beneficiaries by residence status ■ Yearly actual vs planned beneficiaries by age
	COMET, Annual figures: ■ CM-R002b - Annual Beneficiaries by Strategic Outcome, Activity and Modality (CSP): Actual vs planned beneficiaries by strategic outcome, activity and modality	■ Yearly actual vs planned beneficiaries by SO, activity, and transfer modality ■ Yearly actual vs planned beneficiaries by transfer modality ■ Yearly actual vs planned beneficiaries by SO ■ Yearly actual vs planned beneficiaries by activity ***Note**: the number of beneficiaries by strategic outcome, activity and modality includes overlaps if one activity has multiple activity tags.*
	COMET, Annual figures: ■ CM-R020 - Annual adjusted Participants & Beneficiaries by Act Tag, Ben Grp, Gender, Age Grp: Actual beneficiaries by gender, age group, modality (FOOD, CBT and CS)	■ Yearly actual vs planned beneficiaries and participants by SO, activity, activity tag, transfer modality and gender
	COMET, Monthly figures: ■ CM-A003 Actuals - Beneficiaries - Detailed (monthly): Actual beneficiaries by month, activity tag, location, age group and gender	■ Monthly actual vs planned beneficiaries by gender, age group, activity tag and location
	For quarterly figures, MoDa for the ongoing year where ACR is not available: ■ Quarterly Adjusted Estimated beneficiary data: by transfer modality, programme area, residence status, age group and gender	■ Quarterly adjusted estimated beneficiary data by transfer modality, programme area, residence status, age group
Transfers	COMET: ■ CM-R007 – Annual Distribution (CSP): Planned vs actual distribution of food (mt) and cash-based transfers (CBT) (USD) by SO and strategic result. ■ CM-R014 – Food and CBT (corporate report): Annual planned vs actual distribution of food (mt) and CBT (USD) by regional bureau, country office, strategic result, SO, activity and activity tag.	■ Yearly actual vs planned mt and USD distributed by SO and strategic result.

Type	Quantitative data	Possible analysis
VAM, monitoring and evaluation	■ OEV/MIS ■ Publications page[25] ■ Vulnerability analysis and mapping data, including assessments and mVAM data ■ VAM Dataviz[26]: features interactive data visualizations related to food security, climate, and economic indicators. It has further links including the Hunger Hub (a global tool to track food insecurity and key drivers in over 90 countries in close to real-time). Reports explorer presents available reports mainly on food security and market monitoring ■ Audit Reports[27]	■ Overview of monitoring and evaluation exercise ongoing/completed in the country ■ Overview of assessments ongoing/ completed in the country ■ Overview of oversight exercise ongoing/ completed in the country
	■ VAM, Monitoring and Evaluation Planning and Budgeting Tool	■ Analysis of the percentage of CSP budget spent for VAM, monitoring and evaluation ■ Analysis of share of CSP resources spent for VAM, monitoring and evaluation (can be disaggregated by type of exercise for example outcome monitoring, market analysis, evaluation,…) ■ Analysis of the percentage of planned VAM, monitoring and evaluation activities conducted ■ Analysis of staffing (number and grade of people working in the monitoring and evaluation unit and number and grade of people working under each function
	COMET, Outcome indicators data: ■ CM-L008b - CRF Outcome Indicator Values: Activity, location, target group, modality, beneficiary sex	■ Trend analysis ■ Analysis of target achievement
	COMET, Output indicators data: ■ CM-R008 – Output Indicators (CSP): Unit (for example hectares, km, schools, number), target, actual, % achieved	■ Trend analysis ■ Analysis of planned vs actual achievement
	COMET, Cross-cutting indicators data: ■ CM-L009 - CRF Cross-cutting report: Activity, location, target group, modality, beneficiary sex	■ Trend analysis ■ Analysis of target achievement
Gender	Gender and Age Marker: ■ GaM Dashboard, showing the GaM Design codes for current I/CSPs.	■ Analysis of gender integration in the CSP
Supply chain	■ LESS data ■ DOTs	■ Analysis of post-delivery losses
	■ Regional logistics/supply chain key performance indicator (KPI) dashboard	■ Supply chain KPI values

ToC

BY ACTIVITY/ TOPIC		
Type	**Quantitative data**	**Possible analysis**
Asset creation and livelihood support	■ WFP Dashboard it provides information on the number of assets at the global level regional bureau Factsheets	
Climate adaptation and risk management	■ Tableau analytics Climate Risk Management Solutions[28] - Number of beneficiaries (planned and actuals) by climate risk management solution, by country office by year	
Capacity strengthening	■ South-South and Triangular Cooperation (SSTC) Quarterly Newsletter South-South Match	■ Overview of the SSTC initiatives the country is involved in
Emergency preparedness	■ OPweb	
Malnutrition (treatment and prevention)	■ Factory thematic report: CSP NBP, IP and available contributions to Nutrition/ HIV Activities by country office, and year ■ Fill the nutrient gap analysis ■ The cost of the double-burden of malnutrition	
School meals	■ Tableau Analytics School Feeding - Tableau Server (wfp.org)[29] ■ Factory thematic report: CSP available contributions to School Feeding Activities by country office, and year	
Partnerships	COMET ■ CM S007; CM S011 and CM S013	■ Data on Partners' monthly summary of beneficiaries and distribution; CO partners specially on cooperating partner arrangements
	■ DOTS	■ For information on food distribution by cooperating partner and demography of distribution, as reported by them through Partner Connect, when this functionality is available
	■ UN Partner Portal	■ Information on profiles of partners and prospective partners by location, type and review process by United Nations agencies

Acronyms

ACR	Annual country reports
AME	Assessment, monitoring and evaluation
APP	Annual planning process
CBT	Cash-based transfer
CD	Country Director
CEF	Contingency Evaluation Fund
CES	Corporate Evaluation Strategy
CO	Country office
CPP	WFP Corporate and Performance Division
CSP	Country Strategic Plan
CPB	Country Portfolio Budget
DCD	Deputy Country Director
DSC	Direct support costs
DE	Decentralized evaluation
DEQAS	Decentralized Evaluation Quality Assurance System
DEQS	Decentralized Evaluation Outsourced Quality Support Service
DI	Disability inclusion
EC	Evaluation commission
EM	Evaluation manager
EoI	Expression of interest
EPB	Evidence planning and budgeting tool
ER	Evaluation report
ERG	Evaluation reference group
EQAS	Evaluation Quality Assurance System
GEWE	Gender equality and women's empowerment
HQ	Headquarters
ICSP	Interim Country Strategic Plan
IE	Impact Evaluation
IR	Inception report
LTA	Long-term agreement
M&E	Monitoring & evaluation
MCDE	Multi-country decentralized evaluation
MGD	McGovern Dole
MIS	Management information system
MR	Management response
NGO	Non-governmental organization
OEV	Office of Evaluation
PACE	Performance Assessment and Competency Enhancement
PHQA	Post-hoc evaluation quality assessment
QC	Quality checklist
R2	Risk and Recommendation Tracking Tool
RAM	Research, assessment and monitoring
RB	Regional bureau
REC	Regional evaluation committee
RES	Regional evaluation strategy
REO	Regional evaluation officer
REP	Regional evaluation plan
REU	Regional evaluation unit
SDG	Sustainable Development Goal
TDY	Temporary duty
TOR	Terms of reference
TN	Technical Note
UNDSS	United Nations Department of Safety and Security
UN-SWAP	United Nations System-Wide Action Plan
UN-SWAP EPI	United Nations System-Wide Action Plan Evaluation Performance Indicator
VAM	Vulnerability analysis and mapping
WFP	World Food Programme

Endnotes

1 Adapted from UNEG definition in the 2016 UNEG Norms and Standards.

2 Evaluability is the extent to which an intervention can be evaluated in a reliable and credible fashion. Evaluability assessment calls for the early review of a proposed activity to ascertain whether its objectives are adequately defined and its results verifiable. Source: OECD-DAC Glossary of Terms, 2010, p.21.

3 The tool is evolving and is referred to by different names.

4 The structure of the commissioning office may be different for example, no M&E team and only an M&E focal person playing other roles. These roles are to be adapted to the context as appropriate, mindful of the need for impartiality.

5 While most evaluations will focus on programmatic subjects, it is important to involve operation staff including supply chain in these initial conversations.

6 See previous note that this tool is evolving and may be referred to by other names.

7 Experience shows that retaining these costs on the WFP side instead of including in the contractor's budget can reduce the cost by almost half.

8 Between 2021 and 2023, 4 DEs were completed in RBD, only one no-joint DE RBD (Chad country office).

9 This is based on analysis of duration for DEs conducted for DEs starting between 2021 and 2023.

10 The tool should be updated and shared with the regional monitoring adviser and REU in the first quarter of each year.

11 Given WFP work does not necessarily change from CSP to CSP, many of the activities will be continuations. It is therefore important to define what "baseline" really means for each subject of evaluation.

12 The evaluation management group includes representatives from the commissioning agencies that assume evaluation management responsibilities.

13 The steering committee is composed of the heads of the various agencies/entities commissioning the evaluation and it is chaired or co-chaired by one or more of them. The joint evaluation management group is composed of the evaluation managers appointed within each commissioning unit.

14 The IR, addressed in Phase 3, also provides an opportunity to revise the evaluation's expected outputs, overall approach and key lines of enquiries. The ToR and the inception report combined provide the roadmap for the evaluation.

15 There are plans to develop a Technical Note on Evaluability Assessment

16 ToR should focus on clarifying the EQs and the problems they intend to address and challenges envisaged, and outline proposed approach and methodology without being overly prescriptive.

17 As discussed during LTA workshop in October 2023, the Office of Evaluation is exploring options to have a shared file with all DEs that is accessible to all firms in read-only and which REUs and CapQual update regularly.

18 Where there are time pressures and the REU/OEV CapQual focal person considers the draft ToR good enough, the draft can be simultaneously shared with ERG and DEQS which would save thee weeks

19 This has happened especially for JEs with national institutions and can be very damaging if not detected and addressed. Partners have complained directly to WFP CD, which risked becoming a diplomatic issue.

20 Some evaluators new to WFP have misunderstood the requirement to use the template while experienced evaluators have refused to use the templates (or DEQAS more generally) claiming it is too restricting. It is likely they are used to working with agencies with more flexible requirements

21 If under time pressure or the draft IR is considered of sufficiently good quality, this step can be combined with sharing with DEQS which saves around three weeks thus increasing efficiency of evaluation processes

22 wfphotline.ethicspoint.com

23 These standards require to provide disabled employees and members of the public access to information comparable to the access available to others

24 Please note that information on forecasts cannot be published.

25 Please note that this page can be accessed by externals

26 Please note that this page can be accessed by externals

27 Please note that this page can be accessed by externals

28 Please note that information extracted from Tableau server should be validated by the country office

29 Please note that information extracted from Tableau server should be validated by the country office

Photo credits

Cover page: WFP/Sayed Asif Mahmud

Page 2: WFP/Rein Skullerud

Page 4: WFP/Sayed Asif Mahmud

WFP Evaluation

wfp.org/independent-evaluation

wfp.evaluation@wfp.org

@WFP_Evaluation

Via Cesare Giulio Viola 68/70, 00148 Rome, Italy

T +39 06 65131