

Evaluation title	Thematic Evaluation of WFP Philippines Country Capacity Strengthening Activities July 2018 – June 2022	
Evaluation category and type	Decentralized - Thematic	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 67%	
<p>The Thematic Evaluation of WFP Philippines Country Capacity Strengthening Activities July 2018 – June 2022 constitutes a satisfactory report that provides credible findings which decision makers can use with confidence. It clearly summarizes the evaluation purpose, rationale and methodology, and provides information on internal and external contexts. Findings on all evaluation questions are supported by evidence, drawing upon a variety of secondary and primary sources. The evaluation effectively mainstreams gender equality. It formulates conclusions that summarize insights across evaluation findings and criteria and presents five targeted and prioritized recommendations. However, the findings section could have been strengthened by drawing more on the excellent thematic reviews included in the Annexes. The conclusions would have benefited from clearly stating forward looking implications, and the recommendations could have been strengthened by merging thematically overlapping recommendations into one. Readability of the report could have been improved by using more visual aids such as tables, graphs or bullet points, and by omitting some activity-level descriptive details in favour of more analytical synthesis. The description of the methodology would have benefited from elaborating on how the evaluation used the intervention's theory of change and the thematic 'deep-dives' to inform the overall assessment.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The summary accurately and concisely captures key information about the evaluation. It presents key evaluation features and relevant context information, and adequately summarizes key findings related to all evaluation questions. The summary presents an overview of the evaluation conclusions and summarizes all recommendations that appear in the main report. However, the summary exceeds WFP requirements on length.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Highly Satisfactory
<p>The report summarizes relevant aspects of the evaluation context. It strikes a good balance between detail and synthesis, and reflects on gender equality, equity, and inclusion issues. It provides information on how WFP programming in the Philippines evolved during the review period and describes the main features of the evaluation subject in terms of Country Capacity Strengthening (CCS) activities under the CSP, its theory of change, and CCS intervention modalities. The context section would have benefited from including a definition or explanation of how WFP defines CCS, and from including a visual representation of the WFP CCS framework or reconstructed CCS theory of change in the main body of the report.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The evaluation report describes the evaluation objectives, purpose, and scope in terms of the time period and the activities of WFP's work covered. It identifies its intended users and uses, and notes that gender equality and women's rights, as well as accountability to affected populations have been mainstreamed in the evaluation.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The chosen mixed methods approach, and methodology were appropriate for answering the evaluation questions, despite limitations due to the COVID-19 pandemic. Data collection included primary data collection from a variety of stakeholders. One of the evaluation sub-questions explicitly addressed gender equality and wider equity considerations. The use of multiple mutually complementary data sources facilitated triangulation. Evaluation activities were carried out in alignment with relevant ethical standards and took gender, equity and broader inclusion issues into account. The report could have been further strengthened by stating how existing data gaps informed the choice of methodology in</p>		

relation to GEWE; describing more clearly how the evaluation used the reconstructed Theory of Change, and explaining how the 'deep-dive' thematic assessments were to be used to inform evaluation findings and conclusions.

CRITERION 5: FINDINGS	Rating	Satisfactory
------------------------------	---------------	--------------

The evaluation addresses all of the posed evaluation questions and sub-questions in a clear and structured fashion. Evidence is presented transparently and clearly, providing sources for all presented data and quotes, and using a neutral tone. The report discusses WFP contributions to results in a fair and nuanced way, considering contextual factors, including the COVID pandemic. The report reflects the voices of different stakeholder groups from inside and outside of WFP. However, the findings section could have been strengthened by synthesizing some of the very detailed descriptive sections into higher-level observations, and by drawing more, and more explicitly, upon insights deriving from the six thematic reviews. The report could have explicitly noted whether WFP's work had any unanticipated effects, including on human rights and gender equality. Despite the evaluation's intent to assess the intervention against the International Humanitarian Principles, these principles are not referenced in the findings.

CRITERION 6: CONCLUSIONS AND LESSONS	Rating	Partly Satisfactory
---	---------------	---------------------

The conclusions synthesize evaluation findings, noting both strengths and weaknesses of the evaluation subject. The conclusions include explicit reflections on GEWE dimensions, as well as on broader equity and inclusion issues, including disability. However, the conclusions would have benefited from more explicitly formulating strategic implications of the evaluation findings for future WFP programming, e.g., by not merely naming observed weaknesses or gaps but by also elaborating why these gaps (should) matter to WFP. Additionally, several of the presented conclusions do not fully align with, or reflect, previously presented findings, and one of the conclusions introduces new data that had not been mentioned in the findings (related information is presented in the thematic reviews included as Annexes but does not appear in the main report).

CRITERION 7: RECOMMENDATIONS	Rating	Partly Satisfactory
-------------------------------------	---------------	---------------------

The evaluation formulates five prioritized recommendations, one of which addresses GEWE issues. The recommendations include a timeframe for action and identify responsible actors. The recommendations do not consistently and clearly flow from the presented findings and conclusions though and, in some cases, introduce concepts (such as the 'whole of government' CCS approach) not previously discussed in the report. The internal consistency of the recommendations could have been strengthened by merging overlapping recommendations, and by ensuring that all recommendations address issues of similar complexity. The extent to which recommendations are specific and actionable could have been strengthened by identifying main recommendations and related sub-recommendations; or by clarifying how the different recommendations link to, and inform, each other and/or how they could be sequenced.

CRITERION 8: ACCESSIBILITY AND CLARITY	Rating	Partly Satisfactory
---	---------------	---------------------

The report follows the WFP template and includes all of the mandated annexes. It is written in clear and understandable language and makes use of some tables and graphs to convey information, as well as bold font and coloured shading to highlight elements. The report provides sources for all data and quotes and uses cross-references to refer to relevant information entailed in annexes. However, the report is considerably longer than prescribed by current WFP criteria, and would have benefited from summarizing information more concisely, e.g., by using tables. The order in which Annexes are referenced in the main report is not fully consistent with how they are listed. Readability could also have been strengthened by using slightly fewer acronyms and, in some cases, greater language precision to avoid contradictions between findings/conclusion statements and supporting evidence.

Integration of Gender Equality and Women's Empowerment (GEWE) considerations in the evaluation report based on the UN System-Wide Action Plan (UN-SWAP) Evaluation Performance Indicator (EPI) scorecard

UN-SWAP EPI – individual evaluation score	Meets requirements: 7 points
--	------------------------------

GEWE considerations are well integrated into the report. The chosen mixed-method approach and evaluation methodology were based on deliberate considerations on how to effectively integrate GEWE. The evaluation matrix includes one sub-question and several indicators on gender. The evaluation drew upon a variety of data sources and processes, thereby facilitating inclusion, accuracy, and credibility. Findings included reflections on GEWE dimensions, and

two of four recommendations address gender equality issues. Ethical standards were consistently considered, and all stakeholder groups treated with respect for confidentiality and integrity. However, the report would have benefited from stating more clearly to what extent primary data were collected from diverse groups of stakeholders, including both men and women, and representatives from local non-government entities. The report could also have been strengthened by commenting explicitly on potential, positive or negative, unanticipated effects on GEWE and broader equity/inclusion issues deriving from WFP's existing approach to CCS.

Post Hoc Quality Assessment – Rating scale and definitions at overall report and criteria levels	
Highly Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the credible and useful evaluation findings provided and can use the evaluation with a high degree of confidence for decision-making. The report is considered an excellent example.</i>
	<i>Definition at criterion level: The criterion is addressed without any gaps or limitations.</i>
Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making.</i>
	<i>Definition at criterion level: There are no significant gaps or limitations in addressing the criterion.</i>
Partly Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the evaluation findings provided and may use it for decision-making noting that there are some gaps/shortcomings in the information provided.</i>
	<i>Definition at criterion level: There are some significant gaps or limitations in addressing the criterion.</i>
Unsatisfactory	<i>Definition at overall report level: Evaluation users can use some of the learning from the evaluation, noting that there are significant gaps/ shortcomings in the evaluation findings provided. The report may still contribute to decision making but should be used with caution.</i>
	<i>Definition at criterion level: There are critical gaps or limitations in addressing the criterion. Most of the required parameters are not met.</i>