

Evaluation title	Evaluation of Indonesia WFP Country Strategic Plan 2017-2020	
Evaluation category and type	Centralized – Country Strategic Plan	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 89%	
<p>The evaluation of the Indonesia Country Strategic Plan (CSP) 2017-2020 is a strong report, presenting credible findings that users can draw upon with confidence. The analysis of the CSP is clearly structured against key evaluation criteria and questions and is well balanced, supported by diverse sources of data which are well triangulated. The methodology provides a detailed explanation of the evaluation design, data collection and analysis methods as well as the data sources. The linkages among the evaluation criteria, evaluation questions and sub-questions are clear. The report's findings are transparently and impartially generated and effectively present the strengths and weaknesses of the Indonesia CSP, although results data is not systematically disaggregated by age, gender, ethnicity, location, etc. While gender, equity, and inclusion dimensions related to the CSP are well integrated into the evaluation framework, these considerations could have been better reflected in the evaluation conclusions and recommendations.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Highly Satisfactory
<p>The summary provides a concise overview of the evaluation features including on rationale, stakeholders, users, and questions. Graphics and tables are used to present information on country context, overview of evaluation subject, and findings in an effective yet succinct way. The summary accurately reflects the key evaluation findings, conclusions, and recommendations. However, the section would have benefited from the inclusion of a more detailed description of the evaluation purpose and objectives.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory
<p>The context includes a sectoral analysis, information on the internal and external WFP context, and a good presentation of the evolution of Country Programme into the CSP. There is a good description of the CSP, including its objectives, results, activities, and budget with differences from the previous country programme design clearly noted. However, a more robust contextual analysis of vulnerability and climate change issues in the country could have been included, in addition to more information on the intervention logic and key assumptions underlying the CSP's expected results.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Highly Satisfactory
<p>The evaluation's rationale, main users and objectives of learning and accountability are clearly presented. GEWE issues are generally well integrated into the evaluation objectives and the scope is well defined in terms of temporal and programmatic areas covered by the evaluation.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The methodology based on a mixed methods approach and case study design was relevant to answer the evaluation questions. The evaluation framework includes a complete matrix with all of the key components such as indicators and data collection methods and data sources for evaluation questions. Data sources are varied, including documents, people, and sites for observation. Stakeholders participating in qualitative and quantitative data collection represented diverse groups and ethical standards are well covered in annex, including protection of vulnerable respondents during data collection. However, while gender and diversity were included as criteria for sample selection, the report does not include a comprehensive sampling strategy (i.e. sample frame, sample size and rationale) for all data collection methods. Moreover, although limitations to the methodology are described, mitigation measures used to address these should have been presented in the report.</p>		
CRITERION 5: FINDINGS	Rating	Highly Satisfactory
<p>The findings are clearly structured against key evaluation questions and sub-questions and present strengths and weaknesses in a balanced way. They rely on different data sources and types of data and transparently identify any</p>		

gaps in the evidence base. Findings capture not only final outputs achieved and their contribution to outcomes and strategic objectives, but also the evolution of adjustments in strategy, activity, and output delivery in keeping with contextual changes. Unanticipated effects, about the Sulawesi emergency response, capacity strengthening for vulnerability mapping, and greater community engagement in the school meals programme are correctly identified. However, the evaluation findings could have placed more emphasis on the voices or views of different social groups and the International Humanitarian Principles could have been explicitly included in the assessment as they appear to have been relevant in the Indonesian context.

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

The evaluation conclusions are balanced, connect findings across different evaluation criteria and questions and discuss the implications of the findings at a strategic level. They do not include any information not addressed in the findings. However, the conclusions would have benefited from more analysis related to the criterion of coherence. There also could have been a clearer demarcation between internal factors and external factors affecting the results of WFP's work.

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

The recommendations are realistic and actionable. They are prioritized and roles and responsibilities for implementation of the recommendations are clearly indicated. However, they could have been more concise, and more attention could have been given to gender and human rights considerations, particularly with regard to recommendations related to WFP strategic direction, direct engagement, and monitoring and evaluation (M&E).

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

The report is clear, well-structured, and well written. It is adequately sourced, visual aids are effectively used, and key messages are appropriately highlighted in bold. Annexes and information are generally signposted and cross-referenced in the main report.

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 this report. The background section includes details on inequalities related to ethnicity, geographic location, and gender. Sub-questions on gender are integrated in the evaluation matrix, as well as a sub-question to assess the appropriateness of targeting related to geographical areas or population groups. Gender and diversity were included as criteria in sample selection which was purposeful. Data collection instruments included lines of inquiry related to gender equality mainstreaming in the CSP design and implementation. While the findings address the extent to which gender was mainstreamed in the CSP design, there is no assessment of the sufficiency of information collected during the implementation period on specific results indicators to measure progress on human rights and gender equality. Furthermore, it is unclear in the findings how the voices of different social groups were triangulated (by ethnicity, gender, age, etc.) and gender and human rights considerations are not reflected in the recommendations.

Post Hoc Quality Assessment – Rating scale and definitions at overall report and criteria levels

Highly Satisfactory	<i>Definition at overall report level:</i> 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>Definition at criterion level:</i> The criterion is addressed without any gaps or limitations.
Satisfactory	<i>Definition at overall report level:</i> Evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making.
	<i>Definition at criterion level:</i> There are no significant gaps or limitations in addressing the criterion.
Partly Satisfactory	<i>Definition at overall report level:</i> 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>Definition at criterion level:</i> There are some significant gaps or limitations in addressing the criterion.

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>