

Evaluation title	Evaluation of Somalia WFP Country Strategic Plan (2022-2025)	
Evaluation category and type	Centralized - CSPE	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 89%	
<p>The <i>Evaluation of the Somalia WFP Country Strategic Plan (2022–2025)</i> is a credible and reliable report that decision makers can use with confidence. The report is clearly written and demonstrates strong internal consistency. The report provides a well-balanced and relevant overview of the country context, describing the country strategic plan (CSP) as well as key policy, socio-economic, and humanitarian factors that affect WFP's programming in the country. The mixed-methods design is robust and well suited to generating objective and unbiased answers to the evaluation questions. Sequencing of data collection methods, along with factors that informed how data collection tools were designed, are effectively described. Findings draw on a wide range of primary and secondary evidence and present a strong and coherent analysis. Conclusions effectively synthesize these findings and provide useful insights to guide future country strategic plans. Recommendations are forward-looking and actionable and clearly identify the key areas of change envisioned for the next CSP in Somalia. However, in a few instances indicators could have been adjusted to be less descriptive and respond to evaluation sub-questions more adequately. Findings could have provided further disaggregation of primary data given the intended diversity of focus group discussions' participants and more explicitly identified both positive and negative unanticipated effects. Recommendations could have benefited from more differentiated prioritization and sequencing of actions over a longer timeframe.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Highly Satisfactory
<p>The summary evaluation report (SER) presents key information clearly and succinctly, enabling it to serve as a standalone document to inform decision-making. It effectively distills the evaluation features, context, and key messages, and includes all recommendations as presented in the main report. Key messages are clearly organized around relevant themes, and useful figures on the humanitarian context, as well as CSP outputs and outcomes are included. No major gaps or limitations are noted.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Highly Satisfactory
<p>The report provides a well-balanced and relevant overview of the country context as well as the WFP Somalia CSP (2022 - 2025). Its strengths include a clear description of key governance, socio-economic, and humanitarian factors that affect WFP's programmes, as well as an explanation of how gender-specific vulnerabilities intersect with age, disability, and displacement status. The overview of the CSP and its evolution over time is clear, as well as the planned and actual resources by outcome and donor type.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report provides a clear and concise overview of the evaluation's purpose and objectives, including both learning and accountability. In addition, it effectively identifies key stakeholders and users. While the report clearly defines the time period covered by the evaluation, it could have been strengthened by including additional information on the activities and dimensions covered by the evaluation.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>Overall, the methodological design is relevant and appropriate for generating an objective and unbiased assessment against the evaluation questions. The analysis of underlying assumptions and the mapping of CSP activities, outputs, and outcomes within the reconstructed theory of change represent a key strength. Similarly, the tailoring of evaluation sub-questions to the evolving WFP programming context in Somalia, the clear explanation of how data collection methods were sequenced and how data collection tools were designed, and the identification of context-specific</p>		

safeguards to ensure ethical standards are all clearly outlined. However, in a few instances, some indicators could have been adjusted to be less descriptive and respond to evaluation sub-questions more adequately. In addition, the report could have more clearly demonstrated how gender, equity, and inclusion dimensions were considered within the methodology, notably by providing a sex-disaggregation of the focus groups participants.		
CRITERION 5: FINDINGS	Rating	Satisfactory
Overall, the evaluation findings are evidence-based, impartial, and fully address all evaluation questions. Findings are clearly articulated and well-structured, with strong internal consistency across findings. They demonstrate a robust and coherent analysis drawing on a wide range of primary and secondary data and provide a clear explanation of causal linkages, as well as the contextual challenges and enabling factors influencing implementation. However, the findings could have been strengthened by providing further disaggregation of primary data given the diversity of focus group discussions conducted and by more systematically and explicitly identifying both positive and negative unanticipated effects.		
CRITERION 6: CONCLUSIONS AND LESSONS	Rating	Highly Satisfactory
The conclusions are well substantiated and effectively synthesize findings across all criteria and questions. They are presented as clear and higher-level analytical statements that provide valuable insights for the next country strategic plan. In addition, they include an appropriate level of contextual and casual factors as part of the synthesis, with no new information presented.		
CRITERION 7: RECOMMENDATIONS	Rating	Highly Satisfactory
The evaluation's five recommendations are forward-looking, actionable, and identify key areas requiring improvement to support the next CSP. They are logically derived and flow clearly from the findings and conclusions. They effectively identify specific actions that can be taken with responsible and contributing entities clearly identified. However, given their high level of ambition, the recommendations could have benefited from more differentiated prioritization and sequencing of actions over a longer timeframe.		
CRITERION 8: ACCESSIBILITY AND CLARITY	Rating	Satisfactory
Overall, the report is well written, accessible, and clear, demonstrating a logical flow and a high degree of internal consistency. However, in a few minor instances readability could have been enhanced by using footnotes more judiciously and by ensuring that all figures and maps are presented with clarity.		

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	Approaching requirements
The evaluation demonstrates some degree of integration of GEWE, with considerations reflected in the evaluation scope and methodology, as well as an evaluation sub-question. Mixed-methods and ethical safeguards are appropriate, however limitations in the availability and disaggregation of gender-related data constrain the depth of analysis. The conclusions and recommendations appropriately acknowledge GEWE, including a dedicated conclusion (Conclusion 4) and recommendation (Recommendation 3). However, the evaluation could have better integrated GEWE considerations through more systematic disaggregation, fuller analysis of diverse beneficiary perspectives, additional description of methods used to assess GEWE, and inclusion of a sampling frame to indicate the intended breakdown of diverse groups.	
Integration of disability considerations in the evaluation report based on the UN Disability Inclusion Strategy (UN-DIS) scorecard	
UN-DIS individual evaluation score	Meeting requirements

The evaluation terms of reference explicitly calls attention to differential effects on persons with disabilities. Disability inclusion is integrated into the evaluation framework through relevant sub-questions, and persons with disabilities are included in stakeholder mapping and focus group discussions. The findings provide some analysis of disability inclusion, including evidence on targeting and assessments. However, these insights are not carried through to the conclusions or recommendations, limiting the evaluation’s overall responsiveness to disability-related findings.

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:</i> 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>Definition at criterion level:</i> There are critical gaps or limitations in addressing the criterion. Most of the required parameters are not met.