

Evaluation title	Evaluation of capacity strengthening activities to government and local communities in Sao Tome and Principe	
Evaluation category and type	Decentralized – Thematic	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 85%	
<p>The Evaluation of capacity strengthening activities to government and local communities in Sao Tome and Principe presents quality and credible findings that can be used with confidence for decision-making. The report's findings are based on a mix of data collection and analysis methods, including document reviews, interviews, focus group discussions (with a mix of stakeholder categories) and a survey. A strength of the report is the presentation of feedback from stakeholders at all levels, from national-level to the community-level, including school directors, canteen workers, parents and children. However, it would have benefitted from systematically citing documents that are referenced, including documents and data in-text, and from presenting data in a more visual form. Nevertheless, overall, its findings provide a balanced, and nuanced, response to all the evaluation questions. The evaluation's description of the context, the evaluation's rationale and the subject of evaluation are comprehensive but would have benefitted from a summary of the CO's analytical work that informed the intervention's design. The evaluation's conclusions are pitched at a higher level of analytical abstraction, identifying implications of the findings for the future of the intervention. Lessons learned that are presented in the report, however, would have benefitted from a clear formulation, e.g., through the inclusion of introductory or summative statements followed by further elaboration. While most recommendations are logically derived from the findings and conclusion, some would have benefitted from enhanced clarity.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary provides a comprehensive summary of findings across all evaluation questions, which are clearly signposted and presented according to evaluation question. Furthermore, all evaluation conclusions are summarized in a clear manner, and they flow logically from the findings in the summary. The summary also includes information on the scope of the evaluation, and concise information on the context, subject of the evaluation, and methodology. The summary would have benefitted from the inclusion of key information on evaluation features and of a summary of the lessons learned presented in the main report. The recommendations would have benefitted from showing clear linkages to findings, as well as brief information on their prioritization, targeting and timeframe.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory
<p>The report provides a comprehensive overview of the country context, and of the subject of the evaluation. The context section includes information on relevant national policies and strategies, a summary of WFP's work in the country, and includes a detailed description of the COVID-19 situation in the country and its effects. The report provides a concise overview of key information related to the intervention, including its logic and its expected results. The context section of the report misses some specific required sub-elements, however, such as agriculture as part of the Gross Domestic Product, smallholder farmer productivity, data on food insecurity levels, and information on climate change and vulnerability.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Highly Satisfactory
<p>The rationale of the evaluation is clearly outlined, and the report provides detailed information on the main users and stakeholders of the evaluation. The scope of the evaluation is fully defined, in terms of temporal, geographic, and programmatic scope. Gender equality considerations were mainstreamed in the evaluation overall.</p>		
CRITERION 4: METHODOLOGY	Rating	Highly Satisfactory
<p>The evaluation's description of its methodology, both in the main report and its annexes, is comprehensive and detailed, and contains almost all required sub-elements. The evaluability assessment outlines specific evaluability strengths and weaknesses of the intervention. The report provides summary information on data sources, triangulation, sample sizes</p>		

and limitations faced by the evaluation. Overall, these present to the reader an account of how the evaluation methodological design was relevant and able to answer its evaluation questions. However, the report's methodology section would have benefitted from presenting the evaluation questions, and detail of the evaluation's sampling strategy.

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

Overall, the report's findings are transparently and impartially generated without bias. The report presents relevant primary data collected in responding to each question, including survey data (through visuals), as well as quotes from interviews and focus group discussions, without compromising the anonymity of informants. For several sub-questions, there is more than one finding, and this served to address different dimensions within a given sub-question and provide necessary nuance to findings. There is no internal inconsistency among findings. However, the report would have benefitted from systematically citing documents when they are used in the report and from including an analysis of progress of actual versus planned outcomes. Alternatively, the report's evaluability assessment would have benefitted from a clear statement on any evaluability issues that might prevent this.

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

The evaluation's conclusions are pitched at a higher level of analytical abstraction, identify implications of findings on the future of the intervention, and connect findings across different criteria. Conclusions logically flow from the findings and are substantiated by them. They do not include any new information not presented in findings.

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

All recommendations and sub-recommendations are prioritized as either high or medium, are categorized, and are each given specific timeframes for action. Most of them (recommendations 1, 2 and 3) are logically derived from the evaluation's findings and conclusions, and overall, the recommendations are aligned with the evaluation purpose and objectives. Annex 11 of the report provides supplementary information to the recommendations, particularly in providing information that takes into account contextual factors, and specific actions that could be taken in the future. Recommendation 4 could have benefitted from greater clarity.

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

The report is clear, easily understood and utilizes precise and professional language. Sources of quotes are provided (without compromising anonymity of respondents). It also includes brief findings statements, enhancing its readability. While the report meets WFP requirements on length, it does not include page numbers. The report would have benefitted from citing data sources for visuals and tables that draw upon secondary sources of data.

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: 8 points
--	------------------------------

While the evaluation did not have a specific objective dedicated to human rights and gender equality, there is evidence that gender equality considerations were mainstreamed throughout the evaluation. The context section includes a specific sub-section which outlines key aspects of the gender profile of the country. GEWE considerations are reflected in two specific sub-questions in the evaluation matrix. The report includes a detailed description of how the evaluation's gender-sensitive approach to sampling and data collection gave care to ensuring that a diverse range of opinions were heard. The recommendations include one which addresses GEWE issues. While there is a mention of unanticipated effects of the intervention on gender equality in one of the findings. However, the phrasing of this finding could be enhanced for clarity. Conclusions also reflect GEWE-related dimensions of findings for all evaluation criteria.

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.