

Evaluation title	Evaluation of WFP's Support to Smallholder Farmers and its Expanded Portfolio across the Agriculture Value Chain in Bhutan January 2019 to June 2021	
Evaluation category and type	Decentralized - Activity	
Post Hoc Quality Assessment (PHQA) - overall rating	Satisfactory: 86%	
<p>Overall, the evaluation users can rely on the quality and credible evaluation findings and use them with confidence for decision making. The evaluation used a robust methodology that drew on a developmental evaluation design and mixed data collection methods to provide useful insights in a changing context. It consulted multiple stakeholder groups, including WFP staff, national government stakeholders, UN partners, farmer beneficiaries, and school staff. Overall, some 125 individuals were consulted, almost half of which are community beneficiaries. The findings provide answers to all evaluation questions/sub-questions, are supported by a strong evidence base, and are adequately triangulated. Even though the findings are effective at presenting the nuanced perspectives of stakeholder groups –including women – the voices of youth and linguistic minorities could have been presented more clearly. Nevertheless, issues of gender equality and equity are generally well integrated into the findings, conclusions and recommendations. Overall, conclusions are analytical and forward-looking and recommendations propose concrete actions to improve WFP's agricultural portfolio in Bhutan. However, considering the evaluation's aim to generate lessons and good practices, a section on lessons learned could have been included. Similarly, the executive summary could have been improved by adding more detail to the findings section, while making the conclusions briefer.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Highly Satisfactory
<p>The executive summary provides a relevant and concise overview of the evaluation features, conclusions, and recommendations, all of which accurately reflect the content of the main report, including GEWE-related considerations. However, the summary exceeds WFP length requirements.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory
<p>The evaluation report provides a good overview of the country context in Bhutan, including its poverty and food security situation, supported by relevant statistics and also describes national policies in key sectors such as agriculture, nutrition, and education. In addition, the report effectively describes the objectives of WFP's work to improve farmer-school linkages as well as the intervention's logic. However, the context section could have presented in more detail international assistance to Bhutan in the area of agriculture and nutrition. Similarly, the report does not discuss analytical work that might have informed the intervention. The report could have also described the intervention more comprehensively, including its beneficiaries and key partners.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report clearly explains the rationale for undertaking the evaluation at this point, with a clear intent to inform Country Strategic Plan (CSP) annual work plans, provide input into the CSP evaluation in 2022, and inform the development of the CSP 2023-2028. Evaluation users are also clearly identified. However, the report conflates the evaluation purpose and objectives, and the evaluation features present scattered information on the evaluation scope. While the executive summary makes a clearer distinction between the learning and accountability purpose and the evaluation objectives, these specific objectives are not presented in the body of the report.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>Overall, the evaluation is grounded in a robust methodology that draws on mixed methods and a developmental approach that responds to the learning needs of evaluation users in a rapidly evolving context. Data collection methods (desk review, key informant interviews [KIIs] and focus group discussions [FGDs]) and sources are clearly identified. Overall, the interviews involved with a diverse range of stakeholders. In addition, the sampling strategy, data analysis methods and limitations are well outlined and ethical safeguards are clearly explained. These include a complaint and feedback</p>		

mechanisms where evaluation participants could express concerns regarding the evaluation process, which is innovative. While the report briefly mentions an assessment of evaluability, referring the reader to an annex, neither the report nor its annexes clearly reflect an evaluability assessment. Moreover, the evaluation framework could have further reflected key elements from the WFP Gender Policy to guide the gender analysis. Finally, the methodology could have presented more detail regarding the participation of linguistic minorities in data collection.

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

Overall, the findings present a balanced portrait of the intervention's strengths and weaknesses, addressing all evaluation questions/sub-questions transparently and without bias. Findings are generally supported by a strong evidence base, with a clear triangulation process that used multiple methods and sources. The findings are also effective at presenting the nuanced perspectives of multiple groups of stakeholders. Similarly, the evaluation acknowledges when there is insufficient data or when the evidence is inconclusive. In addition, the findings draw on the reconstructed intervention logic to assess WFP's support to outcomes in agriculture, such as increased productivity and income of smallholders, farmer-school linkages and access to markets. Furthermore, the findings present a good gender and equity analysis, which assesses both WFP processes (e.g., WFP capacity for gender) and contribution to gender equality results. Unintended results on gender equality and human rights are also discussed. Although the methodology confirms the participation of youth and representatives from linguistic minorities in data collection, the findings do not clearly convey their views, and the findings do not discuss unintended results beyond those on gender equality.

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

Conclusions logically flow from the findings, present the strengths and weaknesses of WFP's work in the agricultural sector, and discuss implications for the future. Conclusions are analytical and structured around themes, thus spanning the analysis across the evaluation criteria. The conclusions also present a good gender equality and equity-focus analysis, reflecting on opportunities to further women's rural empowerment as well as youth and disability inclusion. Considering the evaluation's objective to generate lessons and good practices, the evaluation could have included a dedicated section on lessons learned from the evaluation.

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

The recommendations flow logically from the evaluation's findings and conclusions and consider the context in which they are to be implemented, including enough detail to make them actionable. Each recommendation is clearly targeted at a specific user, and includes a timeline for action. However, the recommendations could have been more explicitly linked to their corresponding finding(s)/conclusion(s) and should have been clearly prioritized.

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

Overall, the report is well written, easy to understand, and data sources are clearly referenced. In addition, the inclusion of numbered finding statements in bold makes it easy for the reader to understand the main takeaways of the evaluation. Likewise, the evaluation makes effective use of visual aids to convey information in a concise manner. However, both the report and its annexes surpass length requirements. In addition, data collection tools are missing from the annexes, and errors were noted in the report in references to the annexes.

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 is integrated into the evaluation objectives, criteria, and questions. Furthermore, the methodology provides an assessment of the availability of sex-disaggregated data, which informed the data collection plan. The methodology used a mixed-methods approach that drew on a variety of stakeholder groups, including women and men beneficiaries. However, some indicators in the evaluation matrix could have indicated more clearly what the evaluation aimed to examine with respect to GEWE. The findings address gender equality issues in a robust manner and discuss unintended results linked to gender equality. Finally, the conclusions and recommendations integrate GEWE considerations effectively. However, while the context section includes a good gender analysis, the report could have more adequately described the intervention's gendered dimensions.

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.