

Evaluation title	Mid-Term Evaluation of USDA McGovern-Dole International Food for Education and Child Nutrition Program in Nepal 2020 to 2024	
Evaluation category and type	Decentralized - Activity	
Post Hoc Quality Assessment (PHQA) - overall rating	Satisfactory: 75%	
<p>The report constitutes a satisfactory evaluation whose findings can be used with confidence for decision-making. It appropriately summarizes the evaluation's purpose, rationale, methodology, context, and subject characteristics. Drawing on diverse data sources and collection methods, it presents findings on all evaluation questions with supporting evidence, including comparisons with baseline data. Gender considerations are consistently mainstreamed. The report formulates a set of lessons learned as well as overarching conclusions and puts forward four recommendations. It makes good use of visuals like tables and figures. The methodology could have elaborated more on how the evaluation used the programme's theory of change, and the evaluation overall could have benefited from formulating fewer, higher-level evaluation questions. The findings could have been strengthened by better synthesizing activity- and indicator-level evidence into higher-level findings, and by ensuring that findings are consistently grounded in triangulated evidence. The conclusions could have benefited from addressing strategic implications of the findings, and some recommendations could have been more specific and actionable.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary provides a clear synthesis of the evaluation. It captures key evaluation and contextual features, summarizes the main evaluation findings, presents a summary of the evaluation conclusions and includes the evaluation recommendations. The summary exceeds the recommended length and could have benefited from omitting some detail, e.g., by formulating higher-level findings statements and omitting some descriptive detail in the introductory sections. The summary could also have been strengthened by ensuring that it does not introduce new information or details not included in the main evaluation report.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory
<p>The report provides a clear and relevant overview of the evaluation context, and effectively summarizes key information on the subject, including its objectives, activities, beneficiaries, and budget. The context section could have been further strengthened by explicitly commenting on how external events such as the COVID-19 pandemic influenced WFP's and other development partners' work in Nepal, and by deepening the intersectional analysis of social groups affected by discussing dimensions such as caste. The presentation of the evaluation subject could have benefited from elaborating on the project's theory of change/pathways of change, on how and why the current project phase differed from previous phases, and by presenting budget information by project outcome.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report provides a clear overview of the evaluation rationale, objectives, and scope. Gender equality and human rights considerations are mainstreamed in the evaluation's dual objectives of accountability and learning. The report could have benefited from explicitly stating why the evaluation took place at this point in time.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The evaluation's non-experimental design and mixed-methods approach, which used a variety of both qualitative and quantitative data sources and data collection methods, were relevant and appropriate for answering the evaluation questions, including in relation to gender equality and human rights considerations. The evaluation appropriately considered ethical standards and it correctly identifies limitations and related mitigation strategies. The report could have been further strengthened by commenting on availability of monitoring data, including on GEWE-related indicators, and how related data gaps informed the choice of evaluation methodology. Also, it would have been helpful to explain how the evaluation used the project's theory of change to inform data collection and analysis. The description of limitations</p>		

should have provided more detail about what each of them, and about their effects on the evaluation's ability to answer the questions. The evaluation methodology could have benefited from using a smaller number of key evaluation questions to facilitate synthesizing micro-level insights into higher-level findings.

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

The evaluation addresses all the evaluation questions in a structured fashion. Supporting evidence is generally presented transparently and clearly, providing sources for all presented data and quotes and using a neutral tone. The report discusses programme contributions to results, considering contextual factors, and reflects the voices and views of different stakeholder groups. It would have been strengthened by making data triangulation more visible; better balancing between detail and synthesis and formulating higher-level findings; ensuring that all causal claims are supported by evidence; and clearly reflecting the voices of marginalized groups.

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

The conclusions summarize findings across evaluation questions and flow logically from the findings. They include reflections on GEWE and wider inclusion dimensions of the evaluated project and do not introduce new information. The report also presents six lessons, most of which derive from the findings and can inform wider organizational learning. The conclusions could have been strengthened, however, by (i) going beyond a summary of findings to also include reflections on strategic implications; and (ii) placing more emphasis on weaknesses in project design and implementation, thereby establishing a clearer link between findings and recommendations.

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

The report makes four main recommendations that generally align with the evaluation findings. One of the recommendations addresses GEWE dimensions. The recommendations are realistic but could have been more specific, actionable and targeted. In some cases, the categorization of recommendations as either 'operational-short term' or 'strategic-midterm' could have been revised, and the use of priority levels would have been more useful had not all recommendations been considered high priority.

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

The report follows the WFP template for evaluation reports and includes all mandated annexes. It makes good use of visual aids including graphs, tables, and textboxes, and visually highlighted summary findings at the beginning of each sub-section. The report provides sources for all data and quotes and effectively uses cross-references to annexes. Report readability could have been improved by avoiding the use of (statistical) jargon and overly long sentences and paragraphs, as well as by significantly reducing the length of the main report and 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: 8 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 questions and indicators on gender. The evaluation drew upon a variety of data sources and processes, thereby facilitating inclusion, accuracy, and credibility. Findings include reflections on GEWE dimensions, and one of four recommendations addresses gender equality issues. Ethical standards were consistently considered, and all stakeholder groups treated with respect for confidentiality and integrity. The report could have been further strengthened by commenting on the availability of monitoring data on GEWE-relevant indicators.

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.