

<b>Evaluation title</b>	Endline Evaluation of USDA McGovern-Dole International Food for Education and Child Nutrition Programme in Nepal, 2020-2024	
<b>Evaluation category and type</b>	Decentralized - Activity	
<b>Post Hoc Quality Assessment (PHQA) – overall rating</b>	Satisfactory: 74%	
<p>The <i>Endline Evaluation of USDA McGovern-Dole International Food for Education and Child Nutrition Programme in Nepal, 2020-2024</i> is a good quality report that decision-makers can rely on with confidence. A comprehensive examination of sectoral issues along with national policies, plans, and investments provides relevant context for the project and the evaluation. The evaluation rationale, objectives, and scope as well as the primary users and uses of the evaluation are sufficiently addressed. The evaluation uses a mixed methods approach that is robust and draws from a systematic triangulation of diverse quantitative and qualitative data sources. The evaluation questions are tailored to the context and clearly structured against OECD DAC criteria. The findings are generally analytical in nature, clearly structured against evaluation criteria, and address all evaluation questions. They are balanced in tone and grounded in credible evidence gathered from diverse stakeholder groups. The conclusions draw appropriately on multiple findings and are well substantiated by evidence. Recommendations, although broad, are appropriately categorized and prioritized, with an indicative timeframe provided. However, the addition of graphics or visual aids and the further articulation of the evaluation’s rationale and objectives could have strengthened the utility of the executive summary as a standalone document. The evaluation matrix could have benefited from greater specificity and stronger internal coherence linking evaluation questions, sub questions, indicators, and data sources. In a few instances, findings could have been more clearly stated and supported by additional evidence. Some conclusions could have been framed at a higher level of analytical abstraction to allow for further insight on implications for the project going forward. Some recommendations and sub-recommendations could have benefited from greater specificity and further consideration of resource availability. Although broadly forward-looking, more could have been done to clearly set out the change sought from WFP through the recommendations.</p>		
<b>CRITERION 1: REPORT SUMMARY</b>	<b>Rating</b>	<b>Partially Satisfactory</b>
<p>The executive summary provides a relatively concise overview of the evaluation’s findings, conclusions, and recommendations. However, the addition of graphics or visual aids could have enhanced accessibility and readability, and the evaluation’s rationale and objectives could have been more clearly articulated. The line of sight between findings, conclusions, and recommendations in the executive summary is at times unclear or inconsistent and the recommendations presented in the executive summary do not fully align with those in the main report, limiting clarity.</p>		
<b>CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT</b>	<b>Rating</b>	<b>Satisfactory</b>
<p>The context analysis provides a comprehensive examination of sectoral issues relevant to the project (e.g. agriculture, food security, nutrition, education) along with relevant national policies, plans, and investments, and SDG achievement to-date. A thorough description of the project is provided covering strategic objectives, outcomes, activities, planned versus actual budget by activity, and implementation status. Both the context analysis and the project description include discussion of gender considerations. However, the description could have been strengthened by more explicit reference to and explanation of the reconstructed theory of change and the results framework. The context analysis could have been strengthened with greater attention to intersectional aspects of inequality (e.g. geography, disability) and with more systematic presentation of disaggregated data.</p>		
<b>CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE</b>	<b>Rating</b>	<b>Satisfactory</b>
<p>The evaluation report provides a sufficient overview of the evaluation rationale, objectives, and scope as well as the primary users and uses of the evaluation. However greater emphasis on the evaluation’s learning objectives—particularly those related to the transition of districts to government management and accountability to affected populations—could have been strengthened given their prominence in the evaluation terms of reference.</p>		

<b>CRITERION 4: METHODOLOGY</b>	<b>Rating</b>	Satisfactory
<p>The evaluation uses a mixed methods approach with data sources, data collection methods, sampling strategy, and sample selection comprehensively described. The evaluation questions are appropriately tailored to the context and clearly structured against OECD DAC criteria. Ethical risks and safeguarding measures are robustly detailed for each phase of the evaluation. However, the evaluability assessment conducted during the inception phase could have been further reflected in the main report, particularly with regard to implications for data availability and data quality. In addition, the evaluation matrix could have benefited from greater specificity and stronger internal coherence linking questions, sub questions, indicators, and data sources.</p>		
<b>CRITERION 5: FINDINGS</b>	<b>Rating</b>	Satisfactory
<p>The findings are generally analytical in nature, clearly structured against evaluation criteria, and address all evaluation questions, ensuring a comprehensive overview of project performance. They are balanced in tone and based on triangulation of credible data collected from diverse stakeholder categories. Unanticipated results are analyzed and presented. However, in a few instances findings could have been more clearly stated and supported by additional evidence. In addition, findings could have been strengthened through greater disaggregation by sex—particularly in the interpretation of qualitative data collected from direct beneficiaries.</p>		
<b>CRITERION 6: CONCLUSIONS AND LESSONS</b>	<b>Rating</b>	Satisfactory
<p>The conclusions are clearly structured against evaluation criteria and draw appropriately on multiple findings. They logically flow from findings and are well substantiated by the evidence presented. Gender equality and women’s empowerment (GEWE) considerations are explicitly addressed. However, some conclusions could have been provided at a higher level of analytical abstraction, providing insights on implications for the project going forward. In addition, there are a few instances in which claims could have been more fully supported by evidence.</p>		
<b>CRITERION 7: RECOMMENDATIONS</b>	<b>Rating</b>	Satisfactory
<p>There are six broad recommendations provided, each accompanied by several sub recommendations. A clear line of sight is maintained between the findings, conclusions, and recommendations. The recommendations are appropriately categorized and prioritized, with an indicative timeframe provided. However, some recommendations and sub-recommendations could have benefited from greater specificity and further considerations of resource availability and actionability. In addition, although recommendations are broadly forward-looking, more could have been done to clearly set out the change sought from WFP.</p>		
<b>CRITERION 8: ACCESSIBILITY AND CLARITY</b>	<b>Rating</b>	Satisfactory
<p>The report is appropriately structured with a clear and logical flow. Footnotes and signposting are used effectively and tables, graphs, and visual aids appropriately convey key information. Bolding of text effectively highlights key messages. However, the report exceeds word limits and could have benefited from further editing to address conciseness and any outstanding grammatical errors.</p>		

<b>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</b>	
<b>UN-SWAP EPI – individual evaluation score</b>	Approaching requirements
<p>GEWE is mainstreamed across the evaluation objectives, criteria, and questions. The mixed methods design draws on diverse sources with evidence of triangulation. However, the evaluability assessment provides limited information on GEWE related data gaps and the sampling strategy for qualitative data is unclear. While the findings include some disaggregated analysis of quantitative data, the analysis of qualitative data does not provide any insights on differentiated stakeholder perspectives. Some GEWE related unanticipated effects are identified, although additional supporting evidence could have improved credibility and quality. One sub-recommendation explicitly references gender, however further attention on identifying actionable measures to strengthen GEWE in project and government systems could have enhanced the recommendations.</p>	

**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
<p>The evaluation terms of reference give appropriate attention to disability inclusion. Disability considerations are reflected in the evaluation questions and integrated into data collection methods and tools. Evaluation findings contain concrete, although limited, references to disability inclusion, including school support to special needs students and consideration of accessibility in meal provision. However, stakeholder engagement with persons with disabilities or representative organizations is not fully detailed. In addition, conclusions and recommendations do not fully reflect disability-related findings.</p>	

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

<b>Highly Satisfactory</b>	<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.
<b>Satisfactory</b>	<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.
<b>Partly Satisfactory</b>	<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.
<b>Unsatisfactory</b>	<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.