

<b>Evaluation title</b>	Final Evaluation of the McGovern-Dole International Food for Education and Child Nutrition Programme in Guinea-Bissau: From 2016 to 2019	
<b>Evaluation category and type</b>	DE-Activity	
<b>Post Hoc Quality Assessment (PHQA) – overall rating</b>	Satisfactory: 75%	
<p>The Final Evaluation of the McGovern-Dole International Food for Education and Child Nutrition Programme in Guinea-Bissau: From 2016 to 2019 provides credible evaluation findings that can be used by decision-makers with confidence. The evaluation is grounded in a solid methodology that draws on a mixed methods approach to collect data from multiple stakeholder groups. The methodology used to assess impact is particularly robust as it was based on a comparison group to examine key outcomes in WFP-assisted and non-assisted schools. The findings address the evaluation questions and present the strengths and weaknesses of the programme in a balanced manner, as well as positive and negative unintended outcomes. However, findings could have better analysed how WFP activities and outputs contributed to the observed outcomes. In addition, although the findings include a gender analysis and sex-disaggregated data, a more insightful analysis of the programme's contribution to women's economic empowerment could have been presented. While the conclusions provide a good summary of the findings, they do not provide an insightful and forward-looking analysis of WFP's role in school feeding in Guinea-Bissau. This would have been particularly important considering that the McGovern-Dole programme has now ended. Moreover, it is difficult without this information to understand the extent to which the recommendations are feasible, although they flow logically from the findings and are specific and targeted.</p>		
<b>CRITERION 1: REPORT SUMMARY</b>	<b>Rating</b>	Satisfactory
<p>While the executive summary exceeds WFP's length requirement, it provides a clear, accurate, and useful summary of the evaluation. Information on the evaluation type, features, context and subject is presented and key findings, conclusions and recommendations reflect a good overview of those presented in the main report.</p>		
<b>CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT</b>	<b>Rating</b>	Partly Satisfactory
<p>The evaluation report provides a good overview of some key aspects of the country context, including an analysis of the poverty and food insecurity situation in Guinea Bissau. The report also presents an intersectional analysis describing how vulnerable groups are affected by inequalities. However, relevant legal frameworks are only partially described and an overview of development assistance in the country is lacking. In addition, the evaluation subject is not clearly described, with only partial information on geographic coverage and programme activities. Finally, the report does not provide a comprehensive overview of the programme's logic and results chain, nor does it discuss the evolution of the programme over the implementation period.</p>		
<b>CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE</b>	<b>Rating</b>	Satisfactory
<p>The report clearly presents the evaluation objectives of learning and accountability and explains why the evaluation was undertaken at the time that it was. It also identifies the primary users and expected uses of the evaluation. While gender equality considerations are integrated, human rights are not clearly reflected. Moreover, the programmatic scope and the range of beneficiaries covered by the evaluation could have been explained more thoroughly.</p>		
<b>CRITERION 4: METHODOLOGY</b>	<b>Rating</b>	Satisfactory
<p>The report clearly describes the non-experimental design, data collection and analysis methods, data sources, and sampling frame. The use of a control group and econometric methods to assess the impact criterion is well noted. The limitations and mitigation strategies, as well as ethical considerations, are clearly outlined. While the evaluation criteria and evaluation questions are organized in an evaluation matrix, which includes a series of indicators to guide data collection, the sampling strategy does not detail how it sought the inclusion of particularly vulnerable groups. Despite some gaps, GEWE considerations are addressed in the methodology, which was designed to collect sex-disaggregated data. However, the methodology was not informed by an analysis of the programme monitoring and evaluation system, and while the evaluation matrix includes several quantitative indicators to measure questions of effectiveness and sustainability, more qualitative indicators could have been included to explore key hindering and enabling factors.</p>		

<b>CRITERION 5: FINDINGS</b>	<b>Rating</b>	Satisfactory
<p>Overall, findings present a balanced picture of the strengths and weaknesses of the programme and address the evaluation questions. In addition, findings present an interesting discussion on unanticipated results on students and cooks (usually women), and discuss the extent to which the programme addressed recommendations from the midterm evaluation. For the most part, findings integrate a gender analysis and present sex-disaggregated data. Data sources are generally well referenced; however, the report tends to present the views of community stakeholders (parents, teachers, students, etc.) together as opposed to presenting their nuanced perspectives. In addition, some findings — for example on take-home rations (THR) and girls’ attendance — are not substantiated by sufficient evidence. While the evaluation acknowledges the lack of monitoring data constituted a gap to address this question, alternative methodologies could have been proposed to adequately answer this question. Moreover, the assessment of WFP’s contribution to outcomes could have explained more clearly how activities and outputs contributed to expected outcomes.</p>		
<b>CRITERION 6: CONCLUSIONS AND LESSONS</b>	<b>Rating</b>	Partly Satisfactory
<p>The conclusions logically flow from the findings and provide an accurate overview of the strengths and weaknesses of the programme. However, the conclusions could have been more analytical, presenting linkages across criteria, and sufficiently reflecting on the implications of the findings for the future of WFP’s school feeding support in Guinea-Bissau. Such an analysis would have helped the reader better understand the recommendations. Moreover, while the lessons identified generally provide an insightful analysis that could be applied to improve school feeding programming in wider contexts, some lessons could have been further expanded upon. For example, the second lesson notes that the SMC was determinant in convincing teachers not to strike but does not explain why.</p>		
<b>CRITERION 7: RECOMMENDATIONS</b>	<b>Rating</b>	Satisfactory
<p>Recommendations logically derive from the findings and conclusions, are targeted at specific users, and include sufficient detail to make them actionable. The recommendations also adequately address issues of gender equality and equity. However, while the recommendations are prioritized, they are not grouped as either strategic or operational. In addition, given that the evaluation report does not clearly discuss the role of WFP in school feeding in Guinea-Bissau now that the McGovern-Dole programme has ended, the extent to which some recommendations are feasible is unclear.</p>		
<b>CRITERION 8: ACCESSIBILITY AND CLARITY</b>	<b>Rating</b>	Satisfactory
<p>Despite some gaps in its structure, the report is generally well written and uses graphs and tables effectively throughout to convey data concisely, especially from the survey that was conducted in the context of the evaluation. Sections are adequately cross-referenced and data are properly sourced. In addition, key findings are clearly summarized throughout the report. However, some key annexes are missing or incomplete including the timeline for the activity and fieldwork agenda.</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>	Approaches requirements: 6 points
<p>GEWE considerations are integrated into the report to a certain extent. GEWE is mainstreamed across evaluation criteria and several indicators in the evaluation matrix were designed to collect sex-disaggregated data. In addition, the evaluation includes a specific question on the effects of take home rations (THR) on girls’ enrolment and attendance in school. However, the design does not explore issues of women’s economic empowerment and decision-making, which is a missed opportunity considering the programme’s pilot on home grown school feeding. An assessment of the M&amp;E system raises the lack of monitoring data on THR and its effect on girls’ school attendance, but this does not seem to have informed the methodology, which does not propose alternatives to answer the evaluation question on THR. Nevertheless, data collection methods are gender-responsive, with focus group discussions having been conducted with boys and girls separately to gather their nuanced perspectives, although it is unclear whether the sampling frame ensured the inclusion of particularly vulnerable groups. While findings do not present a meaningful discussion of the programme’s unanticipated effects on human rights and gender equality, they do include an analysis of the differential effects of the programme on boys and girls and the report’s recommendations address GEWE.</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.