

Evaluation title	Final Evaluation of the first phase (2015-2021) of the McGovern-Dole Food for Education and Child Nutrition Programme in Côte d'Ivoire	
Evaluation category and type	Decentralized - Activity	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 84%	
<p>The Evaluation of the first phase (2015-2021) of the McGovern-Dole Food for Education and Child Nutrition Program in Côte d'Ivoire provides findings that evaluation users can rely on for decision-making, noting that there are some gaps/shortcomings in the information provided. The report includes a clear account of the evaluation's purpose and scope, along with a presentation of its mixed methodological approach. A rationale for the evaluation approach is clearly provided yet with significant limitations acknowledged by the evaluation team. How these limitations were factored into the analysis and articulation of findings, however, is not clear. Further, without an analysis of the programme theory of change and underlying assumptions, the overall story of change the programme sought to realize was not explored. Instead, findings, conclusions, and recommendations focus on the various "pieces" of the programme, including school access, nutrition, literacy, etc, instead of their interaction as making up a larger effort for desired change. The findings, conclusions and recommendations are useful and logical, based on evidence derived from primary and secondary sources consulted by the evaluation team, including comprehensive surveys. While there is some connection across different evaluation criteria, the conclusions section is largely a summarized reiteration of the findings section. A revisiting of the programme goal and objectives would have helped to articulate more strategic-oriented conclusions based on the findings. Gender equality was mainstreamed throughout the evaluation, however, other equity-related dimensions such as vulnerability and disability are not addressed in the report. Lessons learned are not identified, despite the fact that identifying lessons is stated as part of the purpose of the evaluation.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary presents a clear overview of the evaluand and its context and effectively captures the evaluation's key findings and recommendations. The method for the evaluation is detailed, clearly explaining the approach involved and how analysis was undertaken. The evaluation methodology and conclusions could have been more completely summarized and lessons learned identified in accordance with the evaluation purpose.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Partly Satisfactory
<p>The report includes a thorough account of the programme context, providing the reader with a good understanding of food insecurity in Côte d'Ivoire. The programme is also well described overall, including its objectives, implementation status, various components and activities, budget allotment by phase, WFP's primary partner for the programme, beneficiary numbers, and its geographical coverage. While a logical framework is mentioned, it is not presented nor are the programme's expected results at the output and outcome levels. A presentation of the programme's theory of change and assumptions is not included, nor is a complete account of how all previous evaluation recommendations were addressed by the programme. There is no reference to planned and actual transfers disaggregated by year and by activity.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Highly Satisfactory
<p>The report includes a clear outline of the evaluation purpose, its accountability and learning objectives, and scope. While gender and human rights are not explicitly mentioned as evaluation objectives, gender is clearly mainstreamed in the evaluation's overall objectives. A clearer explanation of gender and human rights as they were approached in the evaluation would have been warranted. The main users and stakeholders of the evaluation are clearly identified, including a wide range of actors at the local, national and international levels. The report defines the time period of the program, the geographical areas assessed, the activities, and stakeholders involved.</p>		

CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The report clearly describes the method and rationale for adopting a quasi-experimental design, along with sampling frame and analysis, to answer many of the evaluation questions. A complex mixed methods design was implemented, allowing comparisons to the baseline data and for demonstrating precise change that had occurred on key areas of importance to the programme and to the evaluation. The methods and sampling frame were diverse, with the use of difference-in-difference (DID) analysis, descriptive statistics, interviews and focus group discussions. Men, women, girls and boys were surveyed within schools and communities. Sub-group analysis exploring gender was featured in the sampling frame. Further, the many limitations to the methods adopted are acknowledged, noting these limitations should be taken into consideration when interpreting the data. Beyond exploring gender, there is no explicit consideration of vulnerability which would have been appropriate. In the evaluation matrix presented in the annex, the data analysis methods simply indicate the type of data, either quantitative or qualitative, but not an approach to analysis as the heading would indicate. Other ethical issues relating to observing and interviewing children, beyond informed consent, should have been addressed.</p>		
CRITERION 5: FINDINGS	Rating	Satisfactory
<p>Findings are presented and discussed in a transparent and unbiased way, and address all of the evaluation questions in a balanced manner. Strengths and weaknesses are noted throughout, and evidence based on both primary and secondary sources is used to back up articulation of findings. The report effectively addresses whether recommendations from the midline evaluation were adopted and explicitly considers the unanticipated effects of the programme, particularly from a gender perspective. The methodological approaches adopted led to multiple findings on a multitude of data points of importance to the programme, yet without discussion and analysis of a theory of change and underlying assumptions, the larger change story and its complexities do not allow for the findings to be better contextualized. Some overuse of direct quotes that are not always adequately sourced was noted in the findings.</p>		
CRITERION 6: CONCLUSIONS AND LESSONS	Rating	Partly Satisfactory
<p>The conclusions represent a summary of the findings, cutting across the evaluation criteria. They clearly flow from the findings and are substantiated. Had the programme objectives been revisited or a theory of change and underlying assumptions been addressed, the conclusions would have been more strategic instead of a summarized reiteration of the findings. Wider equity and inclusion dimensions should also have been addressed in the conclusions while lessons should have been presented given that identifying lessons learned was an explicit objective of the evaluation.</p>		
CRITERION 7: RECOMMENDATIONS	Rating	Highly Satisfactory
<p>The recommendations flow logically from the findings and conclusions, are actionable and feasible and specific to WFP with a specified timeframe. The recommendations, however, do not take into account a broader view of the programme that considers its overall goal and objectives and the interplay between the multiple issues the programme addresses, including improved schooling, access, literacy, gender equality, food security, nutrition and overall health of schoolchildren. A theory of change and examination of underlying assumptions may have helped adopt a broader and more strategic perspective for the recommendations.</p>		
CRITERION 8: ACCESSIBILITY AND CLARITY	Rating	Highly Satisfactory
<p>Overall, the report is well written and follows WFP requirements for structure and content, highlighting key messages in bold throughout. There is reference to information across sections of the report, as well as information provided to orient the reader within sections, and visual aids are effectively used.</p>		

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	Approaches requirements: 6 points
<p>Gender is clearly mainstreamed in the evaluation's overall objectives, featuring strongly in the analysis of available monitoring data and decision-making. The evaluation questions specify an exploration of outcomes based on gender and the methodology used was appropriate to evaluating GEWE issues. Sub-group analysis was used to determine</p>	

whether boys and girls had responded differently to the program, although the DID method is designed to compare changes in treatment schools to changes in comparison schools, not change at the individual level. However, the subgroup analysis by gender did allow determination of whether boys and girls separately see and experience impacts. Sex-disaggregated data was collected and is presented in the findings as unanticipated effects of the programme in relation to gender. Recommendations also address gender considerations. However, human rights are not referenced in the evaluation's primary objectives nor identified as mainstreamed into other evaluation objectives. Beyond exploring gender, there is no definition of vulnerability, which would have been appropriate.

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.