

Evaluation title	Formative Evaluation of WFP Livelihoods Activities in Northeast Nigeria, 2018-2021	
Evaluation category and type	Decentralized	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 76%	
<p>The Formative Evaluation of WFP Livelihoods Activities in Northeast Nigeria 2018-2021 constitutes a satisfactory report that provides credible findings that decision makers can use with confidence. The report summarizes the evaluation purpose, rationale and methodology, and provides relevant context information. Findings on all evaluation questions are supported by evidence, drawing upon a variety of primary and secondary sources. The evaluation effectively mainstreamed gender equality. It formulates conclusions that summarize insights across evaluation findings and criteria and presents six actionable recommendations. While the report meets WFP requirements for length, its readability could have been improved by complementing long narrative paragraphs with more visual aids such as tables, graphs or bullet points, and by formulating visually highlighted explicit findings statements for the main evaluation questions. The descriptive chapters could have been strengthened by providing additional context information and data on core indicators, such as on agriculture or education; and providing a more comprehensive description of the evaluation subject, including by ensuring that related information is placed in the thus-labelled sub-section. The description of the evaluation methodology would have benefited from an explicit discussion of limitations and related mitigation strategies, and by elaborating on how the evaluation used the theory of change to guide data collection and analysis.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Highly Satisfactory
<p>The summary presents key evaluation features and relevant context information, and adequately summarizes the assessment related to all evaluation questions. It presents a concise overview of the evaluation conclusions and summarizes the recommendations. However, not all conclusions are clearly supported by the evaluation findings. Similarly, one of the six recommendations does not clearly derive from the presented findings.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Partly Satisfactory
<p>The report summarizes relevant characteristics of WFP's internal and external contexts, and outlines features of the evaluation subject, including information on how WFP programming in Nigeria has evolved leading up to the review period. It provides a detailed analysis of relevant human rights and gender dimensions. However, the discussion of the external context would have benefited from presenting additional data on basic indicators (e.g., agriculture, education), as well as summarizing, at least briefly, key characteristics of the ongoing armed conflict. It would also have been beneficial had the evaluation separated the discussion of the external context more clearly from the presentation of WFP programming, and presented context information more concisely, e.g., by using bullet points, tables, or graphs. The description of the evaluation subject should have included a summary of the intervention's line of sight or theory of change, and basic information on beneficiary numbers and budget figures. Additionally, the report does not provide information on planned versus actual transfers.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report explains the evaluation rationale and provides an overview of the evaluation's accountability and learning-related objectives, and the evaluation scope, in terms of time period and geographic areas covered. Human rights and gender equality considerations are mainstreamed in the evaluation's objectives. The report could have benefited from providing all information relevant to the evaluation rationale, objectives and scope in one, rather than several, sections.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The evaluation's mixed methods approach was appropriate for the evaluation purpose and enabled the evaluation questions to be answered. The evaluation applied the international evaluation criteria for humanitarian operations and drew upon a variety of qualitative and quantitative data sources. Data collection included participatory elements through</p>		

consultations with a broad variety of stakeholders, including WFP beneficiaries. Gender equality and women's empowerment (GEWE) was effectively mainstreamed using dedicated evaluation questions and indicators, and ethical standards were consistently considered. The report could have benefited from providing an assessment of available monitoring data, including related to GEWE; discussing methodological limitations and mitigation strategies, and explaining how the evaluation used the intervention's theory of change to inform data collection and analysis. It might have benefited from modifying the original evaluation questions to avoid overlap and duplication, and ensuring that all evaluation (sub)questions were covered by relevant indicators.

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

The evaluation addresses all of the evaluation questions in a structured fashion. Evidence is presented transparently, providing sources for all presented data and quotes, and using a neutral tone. The report discusses WFP contributions to results in a fair and nuanced way and reflects the voices of diverse stakeholder groups from both inside and outside of WFP. The report describes both positive and negative unintended consequences of the reviewed activities, including in relation to GEWE. However, it could have drawn more frequently on performance data against programming targets and/or corporate indicators. Findings on effectiveness would have benefited from a more explicit summary and assessment (e.g., in form of a table) of the degree to which outputs and short-term outcomes were achieved, including, if feasible, a comparison of actual versus planned results.

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

The conclusions present both strengths and weaknesses of the reviewed activities across the evaluation questions and criteria. They are substantiated by the findings and, for the most part, logically flow from them. The conclusions extensively reflect on gender equality and wider inclusion dimensions. The report would have benefited from explaining whether and how the sub-headings used to structure the conclusions relate to the intervention's theory of change, and from separating conclusions and recommendations more clearly given that some of the conclusions constitute explicit or implicit recommendations. The conclusions formulated in relation to capacity building and scaling up introduce new information that is not discussed in the findings section. While in other cases, such as the discussion of gender responsiveness, the conclusions provide too much detail and read more as a summary of findings rather than a higher-level reflection. Additionally, the discussion of gender responsiveness is disproportionately long, while other dimensions - such as programming efficiency or coherence - are not or only marginally discussed.

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

The evaluation formulates six actionable and, for the most part, realistic recommendations, including one that specifically addresses GEWE issues and related priorities for action. The recommendations are internally consistent and, with only one exception, logically flow from the findings and conclusions. The recommendations are prioritized, include a clear and realistic timeframe for action, and identify responsible actors. Clarity could have been improved by formulating shorter, more concrete recommendations and selected suggestions for their operationalization, while covering their underlying rationale in the conclusions section. Grouping the recommendations by only one rather than three categories (theme, type and period) might have made the grouping more useful to readers.

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

The report follows the WFP template for decentralized evaluations. It is written in understandable language and makes some use of visual aids, including graphs and selective use of bold font. It provides sources for all data and quotes, and uses cross-references with regard to annexes. The report and its annexes meet WFP requirements for length. However, it could have benefited from a final edit to ensure consistency of style, such as in relation to how to mark references (using footnotes or endnotes) in the text. The evaluation could also have been improved by signposting more frequently when/where relevant information can be found in other parts of the main report, and by using visual aids more frequently to replace the sometimes overly detailed and lengthy narrative paragraphs.

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 are based on deliberate considerations on how to effectively integrate GEWE. The evaluation matrix includes sub-questions on gender. The evaluation drew upon a variety of data sources and processes ensuring that the voices of different stakeholders (men and women, boys and girls, young and older people, people with disabilities) including beneficiaries were heard, thereby facilitating inclusion, accuracy, and credibility. Findings include reflections on GEWE dimensions, and one of six recommendations is dedicated to gender quality issues. However, the report would have benefited from commenting on the availability of WFP monitoring data on GEWE and broader equity and inclusion issues.

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

Highly Satisfactory	<i>Definition at overall report level: 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>
	<i>Definition at criterion level: The criterion is addressed without any gaps or limitations.</i>
Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making.</i>
	<i>Definition at criterion level: There are no significant gaps or limitations in addressing the criterion.</i>
Partly Satisfactory	<i>Definition at overall report level: 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>
	<i>Definition at criterion level: There are some significant gaps or limitations in addressing the criterion.</i>
Unsatisfactory	<i>Definition at overall report level: 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>
	<i>Definition at criterion level: There are critical gaps or limitations in addressing the criterion. Most of the required parameters are not met.</i>