

Evaluation title	WFP Livelihoods and Resilience Activities in Lebanon, 2016 - 2019	
Evaluation category and type	DE – Activity evaluation	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 85%	
<p>The evaluation of the WFP Livelihoods and Resilience Activities in Lebanon, 2016 - 2019 is a high-quality report that meets WFP quality requirements. The evaluation is based on a solid methodological design, draws on diverse data sources and data collection methods, and uses strong data analysis in support of credible findings. Conclusions provide a higher level of analysis, can inform decision-making, and clearly link findings across evaluation criteria. Moreover, recommendations are realistic, actionable, and identify implementing actors, although they do not identify a timeframe for their implementation. While the description of the evaluation design and of methodological limitations is thorough, mitigation strategies are not presented for most limitations identified and the sampling rationale is not explicitly justified for all stakeholders selected to participate in the evaluation.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary effectively summarizes the main features of the evaluation such as its objectives, scope, stakeholders and intended users and uses, and methodology. It also provides a good general overview of the findings and clearly captures the evaluation recommendations. However, the executive summary should have presented a concise overview of all of the conclusions included in the main body of the report.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory
<p>A good overview is presented of the country context, the subject of the evaluation, its coverage, design, main objectives, logic framework, implementation modalities and main partners involved in the intervention along with the intended use of the evaluation. Poverty rates, food security and nutrition levels and other indicators are also presented and, in particular, the overview effectively focuses on the information relevant to the subject of the evaluation, i.e., areas of the country where Syrian refugees concentrate, using latest data available from reliable sources. Planned activities, gender-disaggregated beneficiary numbers and expected results are presented in detail. The overview section provides also a relevant description of the evolution of the intervention over time. However, the overview should have better explained what other international actors contribute to the implementation of other initiatives relevant to the subject of the evaluation in Lebanon as well as other WFP work in the area.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report clearly describes the purpose, objectives, main users and intended use of the evaluation. It also effectively describes gender dimensions explicitly incorporated into the scope of the evaluation and the methodology. In addition, results indicators, the data collected, and instruments used to measure progress with regards to gender equality and human rights are well described. However, the report should have provided a description of the scope of the evaluation specifying its geographic and thematic coverage and explained the rationale for conducting the evaluation at that particular point in time.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The report provides a thorough assessment of available monitoring data and complete description of the evaluation design, data sources, data collection and analysis methods, as well as methodological limitations. The methodology underlines a gender-sensitive approach that was applied throughout the evaluation including the reconstruction of a GEWE-based Theory of Change. The evaluation methods addressed the diversity of stakeholders affected by the intervention, and several consultations with beneficiaries were carried out as part of the data collection phase. Ethical standards adopted, including informed consent, voluntary participation, confidentiality and anonymity, the do no harm principle, are fully described in Annex 9. In addition, triangulation methods were used to ensure the credibility and transparency of the evaluation analysis and findings. However, the evaluation matrix does not present sub-questions</p>		

that would have offered lines of inquiry to further assess specific aspects covered under each evaluation question, and the report should have systematically provided mitigation strategies for all limitations identified. In addition, the sampling rationale for all stakeholders selected to participate in the evaluation should have been explicitly mentioned.

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

The findings are robust and are substantiated by evidence that is correctly sourced and triangulated. Findings demonstrate balance at underlining both strengths and weaknesses of the initiative using neutral language and clearly explain when gaps in information were found and where evidence was inconclusive thus not allowing to fully answer the evaluation question. Furthermore, contributions of WFP activities to strategic outcomes are clearly described and recommendations from previous evaluations are correctly referenced. The report, however, should have provided further details in the discussion and analysis around unintended findings. Finally, the report could have more consistently assessed the performance of the Programme against Humanitarian Principles.

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

The report includes conclusions that are pitched at a higher level of analysis, are forward-looking and correctly link the evidence across criteria and evaluation questions in order to provide an overall assessment that captures the key elements of the evaluation to inform decision-making. Conclusions demonstrate a balance between the evaluation subject's strengths and challenges. Also, conclusions make reference to a good number of GEWE-related aspects. However, an important finding on the appropriateness of the Programme regarding women's engagement and corrective measures to improve this dimension should have been reflected in the conclusions.

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

The report presents recommendations that are realistic and take into consideration WFP internal and external constraints. Recommendations clearly flow from findings and conclusions and are aligned with the evaluation accountability and learning objectives. Importantly, they indicate the specific actors for implementation and provide sufficient details as to how these can take action.. On the other hand, recommendations should have specified a clear timeframe for action and should have indicated whether they are strategic or operational in nature. Only one recommendation addresses GEWE issues.

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

WFP requirements are in general observed regarding the clarity of the information presented, language used, and use of visual aids. Data sources are consistently provided and all cross-references within the report are clearly signposted. Furthermore, all required annexes are included and correctly listed in a separate document. However, the report should have included a list of figures, tables, graphs, and other visual aids. Also, the report could have more systematically used bold to capture key messages, particularly in relation to the evaluation conclusions and recommendations.

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
--	--

Gender issues were treated in the methodology as cross-cutting and were explicitly incorporated into the scope of the evaluation, with associated indicators for most evaluation questions. Although the report presents an assessment of the availability of information regarding progress on gender-related issues, the report notes that the Programme did not systematically analyse outcomes by gender, which also limited the ability of deriving conclusions about the effectiveness of outcomes for men and women. Data collection was designed to consistently gather sex-disaggregated data, but the perspectives of women and girls are not discussed in the report. The context section in Annex 10 could have provided more information on normative or policy elements related to gender equality and other equity issues. Only one recommendation addresses the need for better-defined indicators for each activity or pillar and appropriate disaggregation at least by gender and status (Lebanese vs. Syrian). However, this recommendation does not provide specific information to make it actionable.

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.