

Evaluation title	Évaluation du Paquet Intégré pour la Résilience dans les 20 Communes Prioritaires au Mali de 2018 à 2022	
Evaluation category and type	Decentralized - Activity	
Post Hoc Quality Assessment (PHQA) – overall rating	Highly Satisfactory: 92%	
<p>The “<i>Évaluation du Paquet Intégré pour la Résilience dans les 20 Communes Prioritaires au Mali de 2018 à 2022</i>” is methodologically robust and analytically strong, with a well-articulated rationale, and clear accountability and learning objectives. The report provides a comprehensive and well-referenced analysis of context and programme design, including a clearly articulated theory of change and explicit mainstreaming of gender equality and inclusion across evaluation criteria. The evaluation follows a mixed-methods, theory-based design aligned with OECD/DAC criteria. Findings are analytically rigorous, highly credible, and well triangulated across quantitative and qualitative sources. Conclusions and lessons are coherent, evidence-based, and useful for strategic decision making, and recommendations are concise, forward-looking, actionable, and appropriately prioritized, including a specific focus on transformational gender approaches. The report is well structured, professionally written, and makes effective use of visual aids and signposting. However, the executive summary could have ensured that all summarized findings fully support the recommendations. Conclusions could have more systematically synthesized evidence across evaluation criteria and recommendation feasibility could have been further strengthened by clarifying partner roles in implementation and explicitly addressing resource implications.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary largely functions as a standalone document. It adequately covers the evaluation’s purpose, scope, methodology, and context. Findings, conclusions, and recommendations are generally well summarized and reflect gender equality and inclusion. However, the executive summary could have been further strengthened by including a more comprehensive overview of the evaluation subject and ensuring that the summarized findings and conclusions contain all of the key information needed to support the recommendations.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Highly Satisfactory
<p>The evaluation provides a strong and well-referenced analysis of Mali’s context, clearly linking socio-economic, political, and environmental factors to the Integrated Resilience Programme. It also presents a concise and comprehensive overview of the evaluation subject and a clearly articulated theory of change. The evaluation identifies specific vulnerable groups and provides a brief discussion of national policies and programmes aimed at improving the wellbeing of these groups. While the report includes an analysis of the factors contributing to women’s marginalization, it could have been further strengthened with a deeper intersectional analysis.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Highly Satisfactory
<p>The evaluation clearly explains the evaluation rationale as well as the dual objectives of accountability and learning. The temporal, activity, and geographic scope are identified as well as the key stakeholders and users of the evaluation. Gender equality considerations are mainstreamed across the evaluation purpose, objectives, and scope.</p>		
CRITERION 4: METHODOLOGY	Rating	Highly Satisfactory
<p>The evaluation is underpinned by a strong evaluability assessment that clearly informs the methodological design. Evaluation questions and sub-questions are well aligned with OECD/DAC criteria and appropriate for meeting the evaluation objectives. The report provides an explanation of adaptations from the TOR. The evaluation follows a robust mixed methods theory-based design that enables the evaluation questions to be answered in a well-evidenced and impartial way. The evaluation matrix logically connects all elements in a generally coherent manner. A clear sampling approach and systematic triangulation are described. Methodological limitations are identified, along with mitigation strategies. Gender, equity, inclusion, and ethical standards are well integrated throughout the methodology, data collection, and analysis. However, the evaluation could have been further strengthened by reviewing sub-questions to</p>		

remove repetition and duplications. Moreover, a few indicators included in the evaluation matrix could have been more precise in their formulation.

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

The evaluation presents a strong and analytically rigorous set of findings well supported by high quality data and analysis. The findings demonstrate high credibility, transparently presenting strengths and weaknesses in a balanced manner. The findings include an analysis of WFP's delivery of outputs and the internal and external factors that supported or limited WFP's contributions to expected outcomes. The analysis of how outcomes varied across time, social groups, and geographic regions is a key strength. Gender, equity, and inclusion dimensions are well integrated throughout, supported by systematic disaggregation of data and the inclusion of voices from diverse social groups. However, some key finding statements do not fully reflect all relevant sub-question analyses.

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

The evaluation presents conclusions that are built on a forward-looking analysis that is generally useful for accountability and strategic decision making. The report includes one overall conclusion spanning evaluation criteria and reflecting a high level of analytical abstraction followed by 10 conclusions that correspond to specific evaluation questions. The conclusions are well-substantiated by and flow logically from findings and introduce no new information. Gender equality and women's empowerment (GEWE) dimensions are appropriately reflected, including through a dedicated conclusion and consistent integration across others. The evaluation presents eight lessons that are thematically organized and clearly grounded in the findings and conclusions, offering insights with potential relevance for future interventions. However, the conclusions could have been further strengthened by being more consistently articulated at a higher level, explicitly synthesizing evidence across evaluation criteria rather than primarily structuring them under individual evaluation questions.

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

The evaluation presents a concise and focused set of six forward-looking recommendations that clearly articulate the changes sought from WFP. The recommendations are clearly and logically derived from the evaluation findings and conclusions. They are specific, actionable and prioritized, with clear identification of WFP units responsible for implementation and clearly defined timeframes. GEWE and broader equity and inclusion considerations are well addressed, including a recommendation explicitly focused on promoting a transformational gender approach. However, a small number of sub-recommendations could have benefited from stronger linkages to the findings. In addition, feasibility could have been further strengthened by clarifying partner roles in implementation and by more explicitly addressing resource implications.

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

The report demonstrates a strong logical flow and internal consistency across sections, supported by clear cross-referencing. The writing is clear, professional, and accessible, with appropriate use of footnotes and annexes to present supporting evidence. All data are properly referenced, with effective signposting to relevant sections, and strong use of visual aids to clearly convey key messages.

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
--	---------------------------

The evaluation demonstrates strong integration of gender equality and women's empowerment (GEWE) considerations across the scope, criteria, and analytical framework, including mainstreaming gender across the evaluation purpose, objectives, and scope. The mixed-methods design is well suited to assessing GEWE, with the methodology clearly articulating gender-responsive data collection and analysis strategies, inclusive sampling approaches, systematic triangulation, and adherence to ethical standards. Findings, conclusions, and recommendations consistently reflect a

gender analysis, drawing on disaggregated quantitative data and triangulated perspectives of different social groups. A specific GEWE-related recommendation is provided.

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

The evaluation terms of reference (TOR) explicitly require assessment of gender equality, equity, and inclusion. Disability inclusion is embedded within the evaluation questions, stakeholder mapping, and data collection methods, including the participation of persons with disabilities in focus group discussions. The findings provide disaggregated and differential analysis on persons with disabilities, drawing on household survey data, and disability-related vulnerabilities are reflected in the overall analytical narrative. Conclusions and recommendations appropriately acknowledge multidimensional vulnerabilities and the needs of the most vulnerable groups and zones.

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.