

Evaluation title	Joint Evaluation of the Rome-based Agencies' Resilience Initiative: "Strengthening the resilience of livelihoods in protracted crisis in the Democratic Republic of Congo, Niger, and Somalia" 2017 – 2023	
Evaluation category and type	Decentralized	
Post Hoc Quality Assessment (PHQA) – overall rating	Partly Satisfactory: 56%	
The Joint Evaluation of the Rome-based Agencies' Resilience Initiative (2017-2023) is a partly satisfactory report that provides reliable information while offering opportunities for enhancement. The report effectively presents the intervention context and evaluation's methodology, features, and findings based on diverse data sources. It develops conclusions that span multiple evaluation criteria and offers nine prioritized recommendations. The findings present relevant limitations, including on the extent to which they answer the evaluation (sub-)questions and the presence of inconsistencies and contradictions among them. The conclusions regarding programme strengths and weaknesses should have provided more balance and lessons should have broader applicability. Finally, the report could have articulated the logical connections between findings, conclusions and recommendations in a clearer manner.		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
The executive summary effectively captures key evaluation features and reflects the main findings, conclusions, lessons, and all nine recommendations. The summary could have been strengthened by synthesizing information with less descriptive detail, using more concise language and ensuring fuller alignment with the evaluative assessments in the main report. Additionally, it could have demonstrated clearer connections between recommendations and the underlying findings and conclusions.		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Partly Satisfactory
The report presents information on the context of the three programming countries along with an overview of the evaluation subject's rationale, objectives, activities, donor, partners and budget. The contextual analysis could have been enhanced by incorporating comparable geographic, demographic and assistance indicators across countries and exploring intersecting vulnerabilities among relevant population groups. The evaluation subject description could have been strengthened by providing clearer a description of outcomes and activities. It should have also included sex-disaggregated beneficiary data and more detail on the gender equality and inclusion dimensions of the subject. Finally, it could have also included outcome-based budget breakdowns.		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
The report clearly states the evaluation objectives, rationale, main intended users, and the timeframe covered. The report could have been strengthened by mentioning key evaluation stakeholders beyond the noted primary users, and by clearly stating the evaluation's scope in terms of the geographic and thematic areas.		
CRITERION 4: METHODOLOGY	Rating	Highly Satisfactory
The evaluation's mixed methods design effectively used diverse data sources and collection methods to address the evaluation objectives and questions. It included consultations with a wide range of stakeholders and adhered to UNEG ethical guidelines, while acknowledging data availability constraints and their implications for the approach. The methodological description could have been enhanced by providing more detail on monitoring data availability and quality, and clarifying the framework used for the theory-based approach.		
CRITERION 5: FINDINGS	Rating	Partly Satisfactory
The report effectively presents findings for all evaluation questions and sub-questions, drawing on diverse evidence sources and triangulating different stakeholder perspectives while addressing strengths and weaknesses. The section, however, shows some significant gaps and limitations. Findings should have ensured clearer alignment between the		

evidence presented and evaluation questions by providing more analytical rather than descriptive findings. They could have also maintained consistency among them and elaborated on the specific contributions of individual RBAs to results. They should have also discussed relevant unanticipated effects.

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

The conclusions connect findings across evaluation criteria and questions, identify some forward-looking implications, and reflect GEWE and broader inclusion dimensions. The conclusions could have been strengthened by avoiding restating findings and ensuring that they reflect strengths and weaknesses of the intervention. Furthermore, some conclusions and lessons are not fully supported by the findings. Finally, lessons could have been improved by ensuring wider applicability beyond the evaluated programme.

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

The report makes nine prioritized recommendations that include timeframes for action, and that explicitly address GEWE issues. The recommendations could have been strengthened by ensuring that they clearly flow from the presented findings and conclusions and making each of them specific and actionable. Grouping the recommendations by theme might have been more helpful than labelling each of them by a different subject area. The recommendation on GEWE issues could have been strengthened by also reflecting broader equity and inclusion issues.

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

The report follows the evaluation report template. It effectively uses visual aids such as figures and tables to convey information and uses cross-references. Key findings are highlighted using bold font. The report includes all the required annexes. Its readability, however, could have been strengthened by reducing overall length, using clear and precise language, avoiding long sentences and conducting a final edit to avoid errors such as missing words. The report could have benefited from visually highlighting good practice examples and direct quotes to break up narrative paragraphs. It would also have been helpful for it to include separate sections on "Evaluation Subject" and "Methodology" rather than combining them into one section.

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. GEWE was effectively mainstreamed in the evaluation framework and the evaluation matrix includes dedicated questions and indicators on GEWE. The report includes an assessment of the availability of information on GEWE monitoring data. The evaluation methodology, data collection methods and tools, and approach to data analysis were gender responsive. This involved using a mixed methods approach, conducting consultations with a diverse range of stakeholders, and considering ethical standards throughout the evaluation. The context section includes reflections on intersectional vulnerabilities but could have been strengthened by including observations on intersecting vulnerabilities more frequently, and from including, for example, information on the roles of displacement status, disability, clan/ethnic identity. The report shows mixed performance in triangulating different voices and disaggregating data. It could have benefited from analyzing differences in perspective between women and men in focus group discussions and providing insights into whether and how different social role groups experienced programme benefits differently. The report mentions unanticipated effects on gender equality that derived from programme implementation but could have benefited from explaining what exactly these were. One of nine recommendations explicitly addresses GEWE considerations.

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.