

Evaluation title	Joint evaluation of collaboration among the United Nations Rome-Based Agencies	
Evaluation category and type	Centralized – Strategic Evaluation	
Post Hoc Quality Assessment (PHQA) – overall rating	Highly Satisfactory: 91%	
<p>The joint evaluation of collaboration among the United Nations Rome-Based Agencies constitutes a highly satisfactory report that provides credible findings which decision makers can use with a high degree of confidence. The report concisely summarizes the evaluation purpose, rationale and methodology, as well as the subject of the evaluation and its context. It presents clearly formulated findings on all of the evaluation questions and sub-questions that are supported by evidence deriving from a range of secondary and primary data sources. The evaluation effectively uses the reconstructed overarching theory of change (ToC) for RBA collaboration and a related nested ToC focusing on the notion of 'value added' of this collaboration. The report formulates appropriately high-pitched conclusions that synthesize strategic implications of the evaluation findings, and presents six clear, prioritized, targeted and actionable recommendations. It makes effective use of visual highlights, such as bold font and textboxes, to present information. The report is written in refreshingly direct, yet professional and precise language, which contributes to good readability. However, the evaluation could have reflected its commitment to applying a gender equality lens to data collection and analysis not only in its findings but also in the conclusions and recommendations.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The executive summary provides a concise and useful summary of evaluation features, insights, and summary recommendations. The summary captures key evaluation features and relevant context information, and adequately summarizes key findings related to all evaluation questions. It presents an overview of the evaluation conclusions and includes all recommendations as they appear in the main report. The summary would have benefited from including selected figures from the main report, such as the overarching ToC. Moreover, the conclusions would have been strengthened had they explicitly addressed GEWE, equity and inclusion dimensions.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Highly Satisfactory
<p>The report provides a concise overview of the context of the evaluation subject, including relevant developments in the global, UN and RBA contexts. This includes an analysis of key policies, strategies and other frameworks on gender equality, human rights and inclusion that are relevant to the evaluation subject. Moreover, the report clearly describes the subject of the evaluation by presenting relevant definitions, an overview of RBA collaborative activities carried out during the review period, and a reconstructed overarching ToC for RBA collaboration. However, the overview of the evaluation subject could have benefited from including: (i) information on organizational structures and human and financial resources put in place by the three agencies to manage RBA collaboration; and (ii) explicit reference to findings of previous evaluations.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Highly Satisfactory
<p>The report clearly and concisely presents the evaluation rationale, objectives, time period covered and geographic scope. The evaluation effectively mainstreamed human rights and gender equality considerations in its objectives. The report also clearly identifies the main users and stakeholders of the evaluation.</p>		
CRITERION 4: METHODOLOGY	Rating	Highly Satisfactory
<p>The chosen theory-based mixed-methods approach is appropriate for answering the evaluation questions. Data collection and analysis draw upon a variety of data sources and data collection methods and are guided by a comprehensive evaluation matrix that included sub-questions on gender equality, equity and inclusion. The report mentions the OECD DAC evaluation criteria applied, the quality of available data on the evaluation subject, as well as methodological limitations. The evaluation considered ethical standards and ensured that stakeholders were treated</p>		

with integrity and respect for confidentiality. However, the presentation of limitations could have benefited from a more explicit presentation of mitigation strategies applied or by commenting explicitly where no suitable strategies existed.

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

The report addresses all evaluation questions and sub-questions in a structured way. Findings are consistently supported by evidence, with sources provided for all data and quotes. The evaluation discusses both strengths and weaknesses of RBA collaboration and traces RBA contributions to results. It triangulates the voices of different stakeholder groups inside and outside of the three RBAs. The report notes that the RBAs were able to improve their collaboration on GEWE issues despite the absence of clearly elaborated related objectives or targets. In a few cases, the report could have been more explicit about disaggregating presented data to illustrate differences or consistencies between the views expressed by different (sub-) groups of consulted stakeholders.

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

The conclusions synthesize evaluation findings in relation to all evaluation questions at a higher, strategic level by capturing key characteristics and factors that positively or negatively influence RBA collaboration. The conclusions also address strategic implications of the evaluation findings that can be useful for accountability and strategic decision making. Conclusions reflect both positive and negative findings and logically flow from the findings. However, the conclusions do not explicitly reflect GEWE or wider equity and inclusion dimensions.

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

The evaluation formulates six actionable and realistic recommendations that logically flow from the evaluation findings and conclusions and their strategic implications. While they are prioritized, include a clear and realistic timeframe for action, and identify responsible actors, the recommendations and their sub-elements neither address GEWE issues and priorities for action nor broader equity and inclusion dimensions.

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

The report is aligned with the WFP template for strategic evaluations. It is written in clear and understandable language, which allows it to clearly communicate both strengths and weaknesses of RBA collaboration. The report makes effective use of visual aids including graphs, tables, and selective use of bold font to highlight key findings for all sub-questions. It provides sources for all data and quotes and effectively uses cross-references within the main report and with regard to annexes. Nevertheless, the evaluation could have benefited from presenting slightly fewer yet focused direct quotes from consulted stakeholders, disaggregating such quotes by stakeholder roles and/or sex, and using more textboxes to summarize good practice examples derived from the country case studies.

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
--	--

GEWE is effectively mainstreamed in the evaluation scope of analysis and across the evaluation criteria and objectives. Two evaluation sub-questions addressed GEWE, equity and inclusion dimensions. While the report does not explicitly comment on the availability or quality of data collected to measure progress on human rights and gender equality results, it implies that no such information was collected by the RBAs. The evaluation approach and methodology were gender-responsive and based on deliberate considerations on how to effectively integrate GEWE dimensions in data collection and analysis. This was reflected in the mixed-methods approach chosen, and in the evaluation drawing upon a variety of data sources and processes, thereby facilitating inclusion, accuracy, and credibility. However, survey responses were not disaggregated by sex and while providing such disaggregation may not have added relevant insights in relation to most questions, the evaluation could have looked at responses from men and women separately in relation to questions regarding the RBAC and cross-cutting issues, and in relation to survey and interview findings related to organizational cultures in the three agencies. Ethical standards were consistently considered, and all stakeholder groups treated with respect for confidentiality and integrity. The evaluation context section and findings reflect some degree of gender analysis, although they do not explore what, if any, gender transformative potential might be inherent in RBA collaboration. Findings triangulate the voices of different stakeholders inside and outside of the RBAs. However, none of

the six recommendations or related sub-points explicitly addresses GEWE, equity or broader inclusion issues. While such reference may not have been easily achievable at the level of the main recommendation statements, some GEWE/cross-cutting issue reflections, such as references to SDG 5 on gender equality, could have been included at the level of sub-points provided (e.g. under recommendations 1, 2 and 3).

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.