

Evaluation title	Final Evaluation of Joint Programme 'Enhancing Resilience and Acceleration of SDGs in the Eastern Caribbean' 2020-2022	
Evaluation category and type	Decentralized - Joint	
Post Hoc Quality Assessment (PHQA) - overall rating	Satisfactory: 72%	
<p>Evaluation users can rely on the quality and credible findings of the evaluation of the Joint Programme 'Enhancing Resilience and Acceleration of SDGs in the Eastern Caribbean' 2020-2022 for decision making. Overall, the report is well written and provides easy access to key information, including in the Annex section. The report clearly identifies the evaluation objectives as learning and accountability and is strong in terms of providing the background of the joint programme (JP) and clarifying why the evaluation is taking place at this point in time. The report includes a fair bit of detail on the methodology used and covers the requisite details on ethical considerations and risk and mitigation strategies. The findings cover all of the main evaluation questions and provide a number of concrete examples to reinforce points covered in the sub-questions, although they are not all effectively supported by evidence. The conclusions summarize key findings for each of the six main evaluation questions but could have more effectively connected the findings across the evaluation criteria. While succinct and actionable, the recommendations are essentially process-based in nature. In terms of the integration of gender quality and women's empowerment (GEWE) and social inclusion, which is a critical dimension given the focus on social protection. The report specifically notes vulnerable groups in the Stakeholder Analysis section and occasional references throughout the report. However, it is inconsistent in referencing and analysing the situation for all of the vulnerable groups.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Satisfactory
<p>The summary provides the reader with a general sense of the focus of the evaluation, captures the key findings, lessons learned, and recommendations found in the main report. Although concise, the summary tends to focus more on the positive aspects of the program, overlooking some key challenges. In some cases, this has resulted in a misrepresentation of what is found in the main report. The statement of evaluation purpose lacks some key elements, such as the need for a gender-responsive analysis focused on vulnerable populations. It would have helped had the summary included some basic definitions and unpacked which groups constitute vulnerable groups.</p>		
CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT	Rating	Highly Satisfactory
<p>The report effectively provides a number of background details, including a stakeholder analysis and a list of activities by PUNO, as well as examining previous evaluations undertaken. It also includes a strong focus on the Theory of Change (ToC), including a reconstructed version that sought to address the need to evaluate impacts for a two-year program and to shift the focus to jointness. The report provides useful details on the specific challenges of relevance to social protection for both countries and at the regional level. Tables capturing some of the key issues or topics provide a helpful snapshot of relevant issues to the program. However, some inconsistency was observed in the reconstructed ToC, including a limited focus on the community-level presentation.</p>		
CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE	Rating	Satisfactory
<p>The report clearly identifies the evaluation rationale, as well as its objectives as learning and accountability. The Stakeholder Analysis provides a good analysis and documentation of the primary and secondary stakeholders. However, the scope of the evaluation is not fully clear in terms of the focus at the community level and an explicit objective on gender equality would have contributed to giving more prominence to this dimension in the evaluation.</p>		
CRITERION 4: METHODOLOGY	Rating	Satisfactory
<p>The evaluation framework included the standard OECD-DAC criteria, as well as a standalone criterion on gender, equity, and inclusion, all with corresponding questions and sub-questions. The evaluation included a sound methodological design, mixed data collection methods, and relevant data sources, although the main focus for interviews was members</p>		

of the JP and government partners with a limited focus on beneficiaries which the evaluation team noted was linked to a lack of response. However, the report speaks in generalities about the evaluation team having been able to accommodate gender differences and using participatory methods without providing any further details, particularly as it relates to the participatory methods used.

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

The findings cover all of the main evaluation questions and provide a number of concrete examples to reinforce points covered in the sub-questions. The analysis focuses on the national, regional and community levels. The analysis is particularly strong in examining the positive and negative impacts of the COVID-19 pandemic. The report also explicitly addresses unexpected effects of the program, including on gender and inclusion. However, some of the findings are not effectively backed up by evidence and there is an overemphasis on the achievement of outputs versus of outcomes. While noting that the overall analysis was focused on all agencies of the PUNO, there is limited reference to the specific contributions of WFP. While the analysis related to the evaluation question on gender and equity provides some concrete examples and one reference to a program for families of disabled children, it is weak in terms of systematically considering and noting issues concerning all of the vulnerable groups listed in the stakeholder analysis section.

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

The conclusions summarize key findings in relation to each of the main evaluation questions and the lessons learnt have the potential for broader applicability. However, overall, the conclusions are quite general and do not sufficiently reflect the key findings reported, providing limited information in terms of accountability and strategic decision making. Given that the dual objectives of the evaluation were learning and accountability, it would have been helpful for the analysis to be more overt in making this linkage, particularly as it relates to the accountability dimension. The gender and social inclusion conclusion is particularly weak, citing only one example as evidence that the JP has effectively applied a gender and social inclusion lens to its work.

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

The recommendations are short and quite well written and presented. Given the shift in focus on jointness, they have a strong focus on measures to improve joint programming overall. Responsible actors are identified for implementation of each recommendation. However, several of the recommendations are focused on processes rather than on measures that could improve the overall effectiveness of the programming. Aside from one overly specific recommendation aimed at facilitating access to hardware, consideration of the needs of the beneficiaries appears to have been overlooked. All the recommendations are prioritized as high, which would make it difficult for decision makers to prioritize which ones should be addressed first, although the timing (calendar date) for each gives some indication of the order of priority. The recommendations include only one reference to GEWE and social inclusion, which is general and somewhat difficult to comprehend given that a main finding suggests that the JP has demonstrated results of GEWE and social inclusion at the systemic level.

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

Overall, the report is quite well written and structured as well as clear in terms of referring the reader to annexes that can provide further information. The report follows WFP requirements in terms of length and inclusion of the requisite titles in the annexes. Some minor use of jargon was found in the methodology. The report includes bold in the titles but does not always highlight specific points. Some issues related to the layout in the findings limit the report's readability.

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
--	--

While gender is clearly mainstreamed in the evaluation framework and there is evidence of efforts to collect gender-disaggregated data, the report notes challenges in terms of the availability of such data. The context section notes some of the vulnerable groups per country ranging from the disabled to female-headed households to migrants. The report lists the number of female participants in the evaluation, principally a head count approach to gender analysis, which is

insufficient as it does not capture outcomes such as equitable participation, ability to speak at the training or ability to make use of training. Even reference to the participation of "gender-focused government departments and civil society" assumes that their participation ensures gender responsiveness and places the burden on these representatives as focal points rather than seeking to assess whether all relevant departments involved in the JP are gender responsive. Finally, there is one recommendation and two sub-recommendations related to GEWE.

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.