



Organizational unit: **UNFPA's Sub-regional Office for the Caribbean** Year of report: **2021**

Title of evaluation report: **Evaluation of the 6th Sub-regional Programme of the UNFPA's Sub-regional Office for the Caribbean: Sub-regional Programme Evaluation 2017-2021**

Overall quality of report: **Very good** Date of assessment: **10 December 2021**

Overall comments: The evaluation is very good, evaluating a complex sub-regional programme which operates at the development-humanitarian nexus with different political contexts to consider. While the timeline appeared sufficient, data collection was constrained by time, reaching a limited sample of 55 stakeholders through remote means and excluding end-beneficiaries from the sample. Despite this limitation, the evaluation is well-defined in scope and approach, and utilized a reconstructed theory of change to clearly and comprehensively assess the extent of results achievement. The evaluation was thorough in its analysis of cross-cutting issues and integration of the 'no one left behind' principle, and provided findings and conclusions which were well-balanced between demonstrating achievements and highlighting where there were areas for improvement. This translated to recommendations that were clear, targeted and actionable. The evaluation addressed people with disabilities (PWD) as part of marginalized groups. The UN Disability Inclusion Strategy (2019) was referenced and issues faced by people with disabilities (PWD) were discussed in the Context section. There was not a specific analysis on the extent to which the program is disability inclusive although the report notes there are gaps in collecting data on PWD.

Assessment Levels
Very Good strong, above average, best practice
Good satisfactory, respectable
Fair with some weaknesses, still acceptable
Unsatisfactory weak, does not meet minimal quality standards

Quality Assessment Criteria *Insert assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)*

1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Good
--	----------------------	-------------------	-------------

To ensure the report is comprehensive and user-friendly

1. Is the report structured in a logical way? Is the report easy to read and understand (i.e. written in an accessible language appropriate for the intended audience) with minimal grammatical, spelling or punctuation errors? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Yes	The report is very well-written and easy to understand with very minimal grammatical, spelling or punctuation errors. There are only some areas where font sizes change or text boxes cut off sentences. There is clear distinction between core evaluation elements and chapters.
---	-----	--

2. Is the report of a reasonable length? (maximum pages for the main report, excluding annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations)	No	The report is 86 pages. As a sub-regional programme evaluation, it is expected that this report would be longer than a single country programme evaluation, however, it does exceed the maximum allowable length for even a thematic evaluation. The use of infographics could have helped to streamline this report and the authors could have paraphrased some findings in order to reduce the length of the report and make it more user-friendly.
--	----	---

3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?	Yes	All required annexes are included, however the link included in one annex is not functioning (Annex 7).
--	-----	---

Executive summary

4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) intended audience; iv) Methodology; v) Main results; Vi) Conclusions and Recommendations?	Partial	The summary is clear, presenting the evaluation purpose, objectives, intended audience, methodology, as well as conclusions and recommendations. However, the conclusions are broad and it is not directly clear how they respond to the evaluation objectives, with seemingly little information presented on effectiveness in the different thematic areas UNFPA focuses on. The broad conclusions simply state that outcomes were achieved, but does not unpack any main results which support this conclusion. This would be useful in ensuring sufficient detail is provided to serve as a stand-alone section for evaluation users.
---	---------	---

5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The summary is within the 5 pages maximum length.
--	-----	---

2. Design and Methodology	Yes No Partial	Assessment Level:	Very good
----------------------------------	----------------------	-------------------	------------------

To ensure that the evaluation is put within its context

1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The context is clearly described, providing a situational and gender analysis which emphasizes changes in context as a result of COVID-19 across the thematic areas.
---	-----	--

2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The evaluation report describes the re-constructed Theory of Change and provides a comprehensive assessment of the original programme intervention logic within the findings on effectiveness.
---	-----	--

To ensure a rigorous design and methodology

3. Is the evaluation framework clearly described in the text and in the evaluation matrix? Does the evaluation matrix establish the evaluation questions, assumptions, indicators, data sources and methods for data collection?	Yes	The process for developing the evaluation framework is clearly described. The annexed evaluation matrix includes the evaluation questions, assumptions to be assessed, indicators, as well as the sources and methods for data collection.
---	-----	--

4. Are the tools for data collection described and their choice justified?	Partial	The tools for data collection are described, but the choice not fully justified for each. Mostly, the selection of using a survey for assessing only a few indicators (quality of humanitarian response, quality of interactions, and quality of partnerships) with a limited number of stakeholders targeted (32) is not well justified in terms of adding additional value not found within interviews (same stakeholders in interviews targeted for surveys).
5. Is there a comprehensive stakeholder map? Is the stakeholder consultation process clearly described (in particular, does it include the consultation of key stakeholders on draft recommendations)?	Partial	A comprehensive stakeholder map is not available, though stakeholder consultation is clearly described throughout the methodology chapter, specifically within the methodology for stakeholder selection and overview of data collection tools sections. It's possible that some stakeholder mapping information may have been available in Annex 7 with the 'List of Atlas/JPS projects under the period under evaluation', but the link to the excel worksheet was not functioning. The evaluation report also notes the Evaluation Reference Group debriefing presentation and its refinement of recommendations as a mitigation approach taken against the potential limitation of having a fully participatory approach. The inability to access 'indirect beneficiaries', including marginalized and vulnerable groups was also noted as a limitation that could only be mitigated through consultation with implementing partners who were described as 'representatives' of vulnerable groups.
6. Are the methods for analysis clearly described for all types of data?	Yes	The methods of analysis are described in detail on pages 13 and include thematic and content analysis and contribution analysis. Some limited comparative analysis and descriptive analysis of quantitative data are also used.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Partial	The limitations and mitigation strategies are briefly summarized in Table 5 (page 14) and include notably, the inability to conduct in-person data collection due to COVID-19 and time limitations. However, six months is a standard timeline for many evaluations, so it is not clear why this timeline posed a significant constraint to the methodology unless there were unexpected delays in some phases.
8. Is the sampling strategy described?	Yes	In total, 55 persons were consulted across the thematic areas of youth and adolescents, population dynamics, gender, and sexual and reproductive health. This sampling framework was based on a preliminary assessment of programme data and included the full complement of humanitarian interventions (n=19) and a random selection of development interventions across thematic areas (n=31).
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology enables the collection and analysis of disaggregated data, with 11 males and 44 females included in the sampling frame.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	Limitations to the design included the inability to conduct a comprehensive gender analysis without access to final beneficiaries of UNFPA's intervention areas. However, the evaluation team sought to overcome this by engaging implementing partners working on cross-cutting issues of human rights, disability inclusion and gender equity and by paying attention to these themes within the analysis through specific indicators included in the evaluation matrix.
3. Reliability of Data	Yes No Partial	Assessment Level: Very good
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Interview and focus group discussion insights were regularly cited to support evidence from documents.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation used statistics from other surveys and studies as well as quotes from interviews and focus group discussions as means to support data drawn from programme documents, which are well-cited throughout the report. Limitations to data presented are always noted, when relevant and necessary.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	There is evidence that data has been collected with sensitivity to issues of discrimination. For example, citations made in the report text were anonymous, generally referencing the source but not identifying the specific name. In addition, the evaluation approach was described as gender and human rights responsive. Where data was available on differential perspectives of the impacts of activities on various groups or countries, these were noted.
4. Analysis and Findings	Yes No Partial	Assessment Level: Very good
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	Findings are consistently substantiated by evidence, frequently citing both primary and secondary sources.
2. Is the basis for interpretations carefully described?	Yes	The rationale for the findings is thoroughly explained. The report is explicit in stating when available data is or is not sufficient to assess contribution by UNFPA to the results identified.
3. Is the analysis presented against the evaluation questions?	Yes	The section on findings is organized according to the evaluation questions and indicators.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluation does well in utilizing the reconstructed theory of change to explain cause and effect links. Unintended outcomes were noted more as constraints, such as resource constraints, which affected the ability to reach the most vulnerable populations, such as persons with disability and LGBTQI.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	As noted under sub-criterion 3.3, where data was available on differential perspectives of the impacts of activities on various groups or countries, these were noted. The evaluation covers the Caribbean, and does well to show the differences in outcomes across countries of focus.
6. Is the analysis presented against contextual factors?	Yes	The findings are presented against contextual factors, notably differences in governance structures, COVID-19 and cultural norms influencing progress.

7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The analysis is intentional in its efforts to recognize the diversity of needs of different groups and assess how UNFPA has been able to, or not able to, address them. There is a specific section (4.2.3) which analyzes the extent to which the diversity of needs of adolescents and young people are taken into account.
5. Conclusions		
	Yes No Partial	Assessment Level: Very good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	Conclusions flow clearly from the findings.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Yes	The conclusions provide a thorough understanding of the underlying issues facilitating and hindering achievements. Cross-cutting issues are clearly addressed in relation to beneficiary groups, especially within conclusion two. In addition, issues of human rights and decent work are also considered in conclusions on internal staffing and the impacts of limited resources on staff health and participation in their families.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions are clearly based on evidence from the findings and therefore do not convey bias.
6. Recommendations		
	Yes No Partial	Assessment Level: Very good
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	There are direct links drawn between the conclusions and recommendations.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations are clear and action-oriented, defining the targeted users and proposed timeframe. The human, financial and technical implications are adequately considered within the discussion of operational implications.
3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?	Yes	There is a specific recommendation on 'leaving no one behind', which encourages more active participation of all persons, including the most vulnerable, in programming and decision-making processes.
4. Are the recommendations prioritized?	Yes	Recommendations are prioritized as medium or high, with suggested timeframes which align with the priority level.
7. Gender		
	0 1 2 3 (**)	Assessment Level: Very good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>		
1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?		<p>a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3) 3 - Human rights, gender and the leaving no one behind principle are considered within the specific objectives of the evaluation.</p> <p>b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3) 2 - There was not a standalone criterion on gender and/or human rights, but these themes are specifically considered within coverage and connectedness criterion.</p> <p>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) 3 - There are numerous evaluation questions which integrate GEEW under relevance, effectiveness, sustainability and connectedness.</p> <p>d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?(Score: 0-3) 3- Gaps in data are noted throughout the evaluation and existing gender analyses cited.</p>

<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3) 3 - The evaluation methodology is explicit in the ways it was able to integrate gender considerations and where also there were gaps. Notably, the inability to access final beneficiaries limited the scope of the reports gender analysis.</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) 2 - Both quantitative and qualitative data are available, though the sample size for this evaluation is quite small, especially for the survey.</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) 3 - Triangulation is consistently applied to validate findings.</p> <p>d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3) 2 - The sampling frame does not include consultations with final beneficiaries, and this was well noted as a limitation and mitigated through consultations with end beneficiary representatives (i.e. implementing partners).</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) 2 - Ethical standards are mentioned within the methodology in terms of the specific UNEG Code of Conduct followed and there is no evidence that this was strayed from. Questionnaires include reference to confidentiality as well. However, a section on how the ethical protocols were applied in practice could have strengthened this evaluation's mainstreaming of GEEW.</p>
---	--

<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3) 3- Yes, the background section provides a comprehensive gender analysis with available data, noting contextual changes to key gender and social inclusion indicators as a result of COVID-19.</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3) 2 - The voices of different social groups are not available, though the evaluation report does explicitly note this limitation and highlights the perspectives from different stakeholder groups clearly, even when they differ.</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) 3 - Unanticipated effects on access to services by vulnerable groups is noted as a finding and addressed within conclusions and recommendations as well.</p> <p>d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3) 3 - Specific recommendations addressing GEEW are included.</p>
--	--

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted.
 (**) Scoring uses a four point scale (0-3).

Overall Evaluation Quality Assessment				
Quality assessment criteria (scoring points*)	Assessment Levels (**)			
	Very good	Good	Fair	Unsatisfactory
1. Structure and clarity of reporting, including executive summary (7)		7		
2. Design and methodology (13)	13			
3. Reliability of data (11)	11			
4. Analysis and findings (40)	40			
5. Conclusions (11)	11			
6. Recommendations (11)	11			
7. Integration of gender (7)	7			
Total scoring points	93	7		
Overall assessment level of evaluation report	Very good			
The evaluation integrates adequately cross cutting issues of gender equality, human rights and disability inclusion, even though is not included as part of the evaluation objective.				
Consideration of significant constraints (e.g. COVID-19 or civil unrest)				
The quality of this evaluation report has been hampered by exceptionally difficult circumstances: Yes No				
If yes, please explain:				