

Organizational unit: Sri Lanka CO **Year of report:** 2022

Title of evaluation report: Government of Sri Lanka/UNFPA 9th Country Programme Evaluation (2018-2022)

Overall quality of report: **Very good** **Date of assessment:** 16 May 2022

Overall comments: This is a comprehensive report that provides a thorough overview and assessment of Sri Lanka's 9th Country Programme. Evaluators adopted a mixed-methods approach to collect qualitative as well as quantitative data using document analysis, key informant interviews, and focus group discussions with rights holders, with the sampling process based on detailed stakeholder mapping. The stakeholder map attached in Annex 6 is thorough and organizes stakeholders by CP outputs. The evaluation matrix provided in annex 4 includes evaluation questions by criterion, assumptions to be assessed, indicators, sources of information, and methods for data collection. However, as described in the UNFPA evaluation handbook, it would have been useful to include summarized key findings for each EQ. Methods of analysis are adequately described in the main report. Content analysis was used to identify patterns, common themes, and meaning in the data collected through document analysis, interviews, and FGDs whereas contribution analysis was used to assess how inputs and activities contributed towards outputs and outcomes. The findings are adequately supported by primary as well as secondary/documentary evidence. Key findings are backed by multiple sources of data. The evaluation performs strongly with respect to disability inclusion. Disability is addressed as one of the cross-cutting themes and within two evaluation questions (under Relevance and Effectiveness). It is noted that a focus-group discussion was planned with people with disabilities, however, only one participant showed up due to restrictions related to the COVID-19 pandemic. Similarly, disability is addressed in conclusions, however, it is not explicitly taken up in the recommendations.

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)?	Partial	The report is logically structured with well-defined sections. The writing is clear and easy to understand, however, it would have been useful to use larger font size in graphs and pie charts for increased readability. Rights-based terminology is not used. In future reports, evaluators should be encouraged to use rights based terminology, whereby the 'beneficiaries' of the interventions are acknowledged as 'rights holders' and those with responsibilities for implementing, administering, funding and otherwise supporting the interventions are acknowledged as 'duty bearers'.
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)	Partial	The report is 73 pages (including the Executive Summary), slightly exceeding the maximum length of 70 pages allowed for CPEs.
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	The annexes are complete and include the ToR, list of interviews, evaluation matrix, stakeholder map, and data collection tools.

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?	Yes	The executive summary has all of the essential elements including purpose, objectives, methodology, sampling strategy, intended audience, key results achieved, conclusions, and recommendations.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is useful and provides adequate detail to understand the overall intervention. At 5 pages, it is reasonable in length. The key results, conclusions and recommendations are very succinctly and effectively presented.

2. Design and Methodology	Yes No Partial	Assessment Level:	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	In Chapter 2, the context of the evaluation is thoroughly defined, including development challenges, national strategies, description of issues related to sexual and reproductive health, adolescence and youth, gender equality and women's empowerment, demographic dynamics, and the role of external assistance.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Partial	The Theory of Change (ToC) of the CP is assessed in the report body under the effectiveness criterion, however, it would have been useful to include the actual ToC in the report or in the annexes.

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?	Partial	The evaluation matrix is briefly described in the text and provided in annex 4. It includes evaluation questions by criterion, assumptions to be assessed, indicators, sources of information, and methods for data collection. However, as described in the UNFPA evaluation handbook, it would have been useful to include summarized key findings for each EQ.
4. Are the tools for data collection described and their choice justified?	Yes	The data collection tools used by the evaluation team are described and included documentary review, key informant interviews (KIs), focus group discussions (FGDs).
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)?	Yes	A comprehensive stakeholder map is attached in annex 6 which lists all stakeholder groups by CP outputs. The stakeholder map includes vulnerable groups including youth with disabilities, people with disabilities, older people, and women's organizations. Additionally, it is noted that a debriefing meeting was conducted with CO staff to get feedback on preliminary findings, conclusions and recommendations.
6. Are the methods for analysis clearly described for all types of data?	Yes	Methods of analysis are clearly described. Content analysis was used to identify patterns, common themes and meaning in the data collected through document analysis, interviews, and FGDs; whereas contribution analysis was used to assess how inputs and activities contributed towards outputs and outcomes.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	The methodological limitations as well as their mitigations are described in Table 5. They are mostly related to the COVID-19 pandemic. For example, due to a lack of quality internet services, particularly for rightsholders, the county office arranged in advance for interviewees to have adequate internet access.
8. Is the sampling strategy described?	Yes	The evaluators adopted a purposive sampling strategy to include representatives from different stakeholder groups. These included current and former staff from UNFPA CO, current and former implementing partners, persons living with disability, staff from UN agencies, donors, and rightsholders. Table 4 provides a breakdown of different stakeholders consulted.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	Evaluators used an appropriate methodology to enable the collection of disaggregated data. They presented gender disaggregated numbers of evaluation participants by data collection method, and note under limitations that disaggregated data was not available to identify the benefits to the adolescents and youth, PWD as well as women. To mitigate this limitation, in-depth interviews were conducted with beneficiaries, government representatives, other UN agencies and stakeholders.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	Evaluators adopted an appropriate methodology to assess cross-cutting issues including equity and vulnerability, disability inclusion, gender equality and human right. There was a specific evaluation question on cross cutting issues including gender equality, disability inclusion, and vulnerable groups. Evaluators also note that efforts were made to ensure participation of vulnerable groups in the data collection, however, due to COVID-19 restrictions, this was not achieved to the desired level.
3. Reliability of Data	Yes No Partial	Assessment Level:
To ensure quality of data and robust data collection processes		
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation matrix, which shows various data sources for each evaluation question, clearly demonstrates triangulation.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Data sources are consistently identified and the reliability of sources is briefly discussed in the methodology section. Evaluators note that the team was in constant touch and discussion with the CO staff to clarify issues and contribute to early and continual data validation. Similarly, it is noted that "interviews with CO mainly focused on clarifying and filling gaps on the CP interventions and related information but also validation of the information collected from other sources".
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	Ethical Considerations are clearly described under the methodology section. It is noted that the CPE was guided by UNEG norms and standards of evaluation, Ethical Guidelines for Evaluation, Code of Conduct for Evaluation in the UN System, and Guidance on Integrating Human Rights and Gender Equality (HRGE) in Evaluations, and Guidance on Disability Inclusion in UNFPA Evaluations. Specific considerations included obtaining permission to participate, confidentiality, and ensuring no members of UNFPA staff were present during interviews with external stakeholders. Similarly, the introductory paragraphs in data collection tools make brief note of the ethical considerations followed.
4. Analysis and Findings	Yes No Partial	Assessment Level:
To ensure sound analysis and credible findings		

1. Are the findings substantiated by evidence?	Yes	The findings are adequately supported by primary as well as secondary/documentary evidence. Key findings are backed by multiple sources of data. For example, under the effectiveness criteria, it is noted that "The activities of the National Youth Panel did not proceed mainly due to the COVID-19 pandemic (document review , KIs with IPs, CO and FGD with the youth [2 male, 1 female])". The document sources are then routinely cited in footnotes.
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretations are described where applicable, for example under Partnerships and resource mobilization in findings it is noted that "As discussed, human resources, partnerships and resource mobilization is one of the key areas that the UNFPA CO aimed to and made effort to strengthen. As a result of its very visible efforts in building partnerships and mobilizing resources, the CO got strong support from donors and managed to mobilise an additional US\$3.63".
3. Is the analysis presented against the evaluation questions?	Yes	The evaluators have organized the analysis by evaluation criteria and associated questions. At the beginning of each criteria findings there are also summarized key findings and challenges.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Under Effectiveness, causal linkages between outputs and outcomes are articulated in the achievement tables for each strategic plan outcome area and then further explained in the subsequent text. EQ 5 focuses on unforeseen consequences; it is noted that there were no unforeseen consequences for any of the outcomes.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	Where possible, evaluators show differential outcomes for the various target groups. For example, under the effectiveness criterion, it is noted that " The target groups again differ in emphasis depending on the province but generally include poor families, marginalized and vulnerable groups, key populations, PWD, ex-combatants, women headed households, orphans, estate workers. Those most affected by war emphasize peacebuilding and dealing with marginalized groups affected by war including ex-combatants (document review)".
6. Is the analysis presented against contextual factors?	Yes	The findings are clearly shown against contextual factors including COVID-19 and socio-cultural context. For example, under the effectiveness criteria, it is noted that "the CO supported interventions that provided platforms for youth participation, and enhanced their capacity for policy dialogue. It also supported the achievement of increased capacity for the implementation of CSE in and out of school. However, progress has been slow due to religious, political and cultural challenges".
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Findings (particularly under relevance and effectiveness criteria) provide analysis on cross-cutting issues including vulnerability, disability inclusion, gender equality, and human rights. For example, under section 4.3.1 it is noted that "UNFPA, with UN Women and UN Volunteers promoted and enhanced the capacity of youth and women to participate in the peace building process. This provided women and youth from diverse populations (ex-combatants, youth from poor socio-economic backgrounds, war widows, people with disabilities, school children, youth from different ethnic groups) the opportunity to engage in peace building and develop capacity in policy development".
5. Conclusions	Yes No Partial	Assessment Level: Very good
To assess the validity of conclusions		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions are clearly derived from the findings; they specify the question(s) and criteria upon which each was founded.
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 are successfully crafted at a higher level than findings. Each has a succinct overall statement which is then briefly expanded upon. They adequately address cross-cutting issues such as gender, vulnerability, aged, and people with disabilities. For example, conclusion # 2 highlights that "CP integrated cross cutting issues of gender equality, disability inclusion and human rights-based approaches. However, the cultural, religious and political environment presents significant risks to the achievement of UNFPA results".
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	Given the conclusions' direct link to findings, there is no indication of bias.
6. Recommendations	Yes No Partial	Assessment Level: Good
To ensure the usefulness and clarity of recommendations		
1. Do recommendations flow logically from conclusions?	Yes	As was done with conclusions, the recommendations are presented by strategic and programme level. The recommendations are based on findings and conclusions, and they show the linkages to specific conclusions through numbering.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	There are a total of 13 recommendations clearly directed toward intended users. However, some of the recommendations could offer more directions on their technical and financial implications of implementation such as recommendation #6 and #9.
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	The recommendations appear to be balanced and impartial.
4. Are the recommendations prioritized?	Partial	The recommendations are prioritized, however, they are all prioritized as 'high' and do not mention the time-frame for their implementation.

7. Gender	0 1 2 3 ^(**)	Assessment Level:	Very good
To assess the integration of Gender Equality and Empowerment of Women (GEEW) ^(*)			
1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?		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) This was not evident in the objectives, evaluation purpose, or scope. (0) 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) Gender and human rights were mainstreamed into the criteria including in relevance, effectiveness, and sustainability through the evaluation questions. (3) c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) There are dedicated evaluation questions (EQs), assumptions, and indicators addressing issues related to GEEW. (3) 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) The evaluation matrix included in Annex 2 has indicators focusing on human rights and gender equality. (3)	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		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) Evaluators note that the evaluation was guided by guidance on Integrating Human Rights and Gender Equality (HRGE) in Evaluations. Under the methodology section, it is noted that an all-female FGD was also conducted with GBV shelters' trainers. Additionally, the list of evaluation participants is disaggregated by gender. (3) 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) The evaluation methodology used mixed methods appropriate for evaluating GEEW considerations. A total of (36 female and 30 male) individuals participated in data collection. FGD participant groups include Youth Panel, Senior Citizens Committees, and GBV shelter counsellors. (3) c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Data was collected from a diverse range of data sources and there is a description provided of processes to ensure accuracy and validity such as getting feedback from ERG and CO on preliminary findings and cross-checking findings through interviews with CO staff. (3) 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) Participants in the evaluation included members of marginalized and vulnerable groups such as the Youth Peace Panel (2 females, 8 males), Senior Citizens Committee members (7 females, 11 males), and GBV shelter trainers (6 females, all females) who had participated in the program to capture their experiences. However, it is noted that, based on the available data, it was not possible to determine the extent to which the most vulnerable populations were involved. (3) e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) Ethical considerations including their application are clearly described. (3)	
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		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) The country context section provides a solid description on issues related to human rights, gender equality, people with HIV, people with disabilities, and SRHR. (3) 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) Voices and perspectives of marginalized and vulnerable groups were captured by conducting FGDS with Youth Peace Panel, members of the Senior Citizens, and GBV shelters' trainers. These are also reflected in the analysis within the findings section. EQ4 particularly focuses on the extent to which vulnerable and marginalized groups (such as young women and girls, persons with disabilities, etc.) had access to life-saving services during humanitarian situations (3) c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) It appears that efforts were made to consider these and that none were found. Under the methodology section it is noted that "the CPE identifies key unintended effects", and then under the relevance section of findings it is mentioned that there were no unforeseen outcomes (2) 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) Several recommendations cover issues related to GEEW. Recommendation # 10 particularly focuses on gender equality and women's empowerment - "UNFPA should support efforts to fully exploit the data and knowledge generated, and the capacities built to further the GBV / IPV agenda". (3)	
(*) 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			
	Assessment Levels (*)		
Quality assessment criteria (scoring points*)	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	69	31		
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.				
The cross-cutting themes, including disability inclusion, are adequately addressed.				
Consideration of significant constraints (e.g. COVID-19 or civil unrest)				
The quality of this evaluation report has been hampered by exceptionally difficult circumstances: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No				
If yes, please explain:				