

Organizational unit: UNFPA Timor Leste Country Office

Year of report: 2020

Title of evaluation report: Evaluation of the 3rd Country Programme 2015-2019/20 in Timor-Leste

Overall quality of report: Good

Date of assessment: May-20

Overall comments: Overall, the evaluation report presents a clear analysis of key findings, drawing on both qualitative and quantitative data from a diverse group of stakeholders and providing a comprehensive description of contextual factors impacting results achievement. Particular strengths of the evaluation are found within the conclusions and recommendations, which are clear and actionable. The methodology could be improved through a clear stakeholder analysis, description of analysis methods and ethical practices, as well as a reconstructed Theory of Change. While this was suggested in the ToR, it was not required, and the evaluation opted out of using a ToC without a clear explanation. The presentation of findings, including the presentation of cause-and-effect links and unintended outcomes, could have been clarified through the use of a ToC. GEEW considerations are adequately integrated into the design, methodology and presentation of findings, though there is no discussion of how the methodology was gender responsive, and as the tools are not provided it is not possible to see how qualitative data were captured. The report draws heavily on the document review, which is well cited, though could be improved through more direct reflection and disaggregation of gender data and quotes drawn from interviews and focus groups, as despite challenges in data collection, a diverse group of stakeholders were reached.

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:

Fair

To ensure the report is comprehensive and user-friendly

1. 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?

Partial

The report is mostly well-written with few spelling and grammatical errors, which do not significantly detract from the overall clarity of the report. Some content within the methodology section is directly copied from the ToR, which is not best practice, and the paragraph spacing varies across sections. As such, the clarity of the reporting provided through the structure could be improved.

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 maximum length for a CPE is exceeded by 6 pages.

3. Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?

Yes

The regular structure is followed.

4. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological tools used (e.g. interview guides; focus group notes, outline of surveys) as well as information on the stakeholder consultation process?

Partial

Missing from the minimum requirements are the methodological tools used. Although there is not a specific section on the stakeholder consultation process, the list of stakeholders consulted is labeled as a Mapping of Consulted Partners and clearly shows the range of groups involved. It should also be noted that the evaluation matrix is very thorough and includes findings data.

Executive summary

5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?	Partial	The conclusions are many and lack sufficient detail to provide a clear summary of the main results of the evaluation, limiting the Executive Summary's usability as a stand-alone section.
6. Is there a clear structure of the executive summary, (i.e. i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?	Yes	The recommended structure is followed, including intended audience.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	It is concisely presented in 4 pages.
2. Design and Methodology	Yes No Partial	Assessment Level: Fair
<i>To ensure that the evaluation is put within its context</i>		
1. Does the evaluation describe the target audience for the evaluation?	Yes	There is a subsection, Evaluation Audience, that lists the range of groups the evaluation results will be shared with.
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	There are comprehensive and well-referenced chapters on Country Context and UNFPA / UN Response & CP Strategies that cover constraints faced.
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	This is difficult to rate. The ToR specifies that the evaluators are to assess the ToC and reconstruct a new one if the existing one is found insufficient. However, under 1.2.2 Evaluation Approach (p 2) it is stated that the "evaluation was not theory-based as the CO requested for the CPE team to focus on conducting interviews and other CPE tasks rather than on reconstruct (sic) the Theory of Change for this CP". Presumably this determination was made at the Inception stage but this change could have been made more explicit in the report. The evaluators do provide an assessment of the ToC; this is done under Efficiency (p 55-56).
<i>To ensure a rigorous design and methodology</i>		
4. 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 evaluation matrix (annex 5) is comprehensive and includes all required elements plus findings for each question. It is referenced in the main report.
5. Are the tools for data collection described and their choice justified?	Yes	Data tools include document review, KIIs, FGDs, a questionnaire and site visits. A succinct explanation for each is provided. The description of document review is notable for specifying the types of documents assessed (compared to many reports that only use general terms such as 'programme documents').
6. 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	Stakeholder mapping is a requirement of the ToR. As mentioned above, the annex includes a list of stakeholders involved in the evaluation and it is helpfully organized by stakeholder type; however it is not clear if this list constitutes all of the entities with which UNFPA works, and therefore cannot be determined to be a comprehensive map. The evaluators emphasize that a participatory process was used and that preliminary evaluation results were shared with the ERG and main CP stakeholder. It is clarified under Evaluation Phases that key stakeholders were part of a validation workshop in which preliminary findings and recommendations were presented.
7. Are the methods for analysis clearly described for all types of data?	Partial	There is one paragraph on data analysis. It discusses analysis quite generally - "a systematic organisation, comparison and synthesis process was undertaken". It does not describe how the different types of data were assessed, but mentions them in terms of the results being used for triangulation purposes for each evaluation question. A more detailed account of how each type of data was analysed would have been useful.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	Noted limitations include the limited focus on programme outcomes/results for beneficiaries, data not being statistically representative, non-availability of some programme documents as well as stakeholders during the time of field visits, and issues with accessibility to field sites. Mitigation strategies were identified as triangulation of information and validation of findings with CO staff and stakeholders.

9. Is the sampling strategy described?	Yes	The report states that a purposive sampling approach was used and provides the parameters for the selection for each of six stakeholder groups. The rationale for the selection of site visits (to 3 municipalities where UNFPA support was the largest) is clearly stated.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The methodology section includes a table of stakeholders interviewed; it is disaggregated by stakeholder group and gender. Otherwise, there is minimal emphasis on, or presentation of, disaggregated data in the report.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The design appears appropriate for this purpose. The respondents included a range of beneficiaries, and the evaluators were purposeful in seeking input from female and male beneficiaries and representatives of partner implementing groups.
3. Reliability of Data	Yes No Partial	Assessment Level: Good
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluators emphasize the use of triangulation to ensure data validity. The triangulation process included the validation of preliminary findings with key stakeholders.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Data sources are very clear in most sections. The evaluation approach relied heavily on document sources and on qualitative data from KIIs. The extensive document citations, the number (106) of interviews conducted, and the range of groups consulted suggests that reliable data were used.
3. Did the evaluation make explicit any possible limitations (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues?	Yes	As mentioned above, data limitations in both primary and secondary data sources were mentioned. The evaluators were clear that the purposive sampling strategy did not lead to perspectives being obtained from a representative sample of beneficiaries, but that the results obtained were indicative. The main mitigation strategies were triangulation and validation; and these appear to have been appropriate for this evaluation.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	A good range of stakeholders, including beneficiaries of different activities, were consulted. However, there is no explicit description of ethical practices followed or referenced to use of UNFPA or UNEG guidance on ethical considerations. Since the evaluation tools were not attached, it was not possible to see if issues such as confidentiality were addressed in the protocol instructions or introductions.
4. Analysis and Findings	Yes No Partial	Assessment Level: Good
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	In most cases the evaluators carefully back up findings with evidence from documents, KIIs or observations.
2. Is the basis for interpretations carefully described?	Yes	The evaluators use highlighted text boxes to summarize the main findings. The subsequent text then shows the basis for these in considerable detail.
3. Is the analysis presented against the evaluation questions?	Yes	The evaluation questions are listed at the beginning of each criteria section and the analysis is structured accordingly.
4. Is the analysis transparent about the sources and quality of data?	Yes	Document sources are cited throughout (through footnoted document references) and there is reference to perspectives obtained during the evaluation from stakeholders. The evaluators draw most heavily on document sources, particularly under Relevance and Effectiveness, while the findings for Partnership, Coordination and Sustainability draw more from KIIs. The evaluators note where lack of data interfered with findings being drawn.

5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Partial	The causal linkages are mostly evident throughout the Findings, though could be made more clear. While the progress for each CP Strategic Result is clearly presented against the indicators and targets in table format and then further elaborated upon in the subsequent text, there is not always a clear causal link drawn between the outputs and outcomes, or analysis of where the chain of logic could be strengthened. Furthermore, while the evaluation suggests that no unintended outcomes were found, it does appear that some of the results achieved are not captured in the CP indicators. The reconstruction of the Theory of Change suggested in the ToR could have helped improved clarity around both cause-and-effect links and unintended effects. However, the reasons for achievement and non-achievement of targets are explained.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	Much of the analysis is focused on output-level results. The evaluators show how the CP responded to meet the needs of marginalized groups as these needs became more apparent and as opportunities emerged. For example, based on a request by Female Parliamentarians, UNFPA helped to fund a study on teenage pregnancy and early marriage and the results were found useful for lobbying government on the importance of vulnerable groups having access to youth activities. There is also some outcome-level analysis - for example, the evaluators note that UNFPA's work to ensure that people with disabilities had access to GBV services led to national health sector GBV guidelines being updated to integrate PWD as a target group.
7. Is the analysis presented against contextual factors?	Yes	Findings are analyzed against relevant contextual factors. Constraints are clearly articulated under Relevance - subsection 4.1.3 Main External Challenges. The Efficiency section also performs well in providing the context for shifts in financial resources and how UNFPA responded to budget cuts.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	This is seen most clearly under Effectiveness where, for example, the evaluators discuss how the CP started supporting HIV-prevention awareness raising amongst key vulnerable populations in border areas (p 34). The subsection on Equality and Human Rights Principles then succinctly provides further analysis of cross-cutting issues.
5. Conclusions	Yes No Partial	Assessment Level: Good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	This section is well structured to show the links between each conclusion and the corresponding findings (noting the evaluation question) and associated recommendations.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Partial	The conclusions effectively capture the findings for each criteria but in some cases present the underlying rationale in more detail (by re-stating findings) than necessary.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	Bias is not apparent in the conclusions.
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	This is clearly done with the corresponding conclusion number(s) given for each recommendations.
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations are clearly presented and include the targeted users and operational implications.

3. Do recommendations appear balanced and impartial?	Yes	They appear balanced by focusing on work that should continue and areas for improvement. They appear impartial.
4. Is a timeframe for implementation proposed?	Yes	The introduction to this section notes that the recommendations are intended for the final year of the current CP and for the next CP.
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	They are given a priority rating (all but one are High) and are framed appropriately for a management response.
7. Gender	0 1 2 3 (**)	Assessment Level: 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) HRGE is not explicitly highlighted in either the evaluation objectives or scope = 0</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) GEEW is mainstreamed and explicitly addressed under Effectiveness and Sustainability. = 3</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) There are 2 GEEW-related questions. EQ5 assesses the extent that the CP outputs contributed to planned GEEW outcomes. The evaluation team added this question to the list set out in the ToR. And EQ8 looks at how UNFPA support has helped ensure gender and other equity issues have been appropriately integrated into national-level policies and instruments = 3</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) This was done and flows through to a recommendation on the need to continue to support population data systems so that equity issues can be better addressed. = 3</p>
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		<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) There is no discussion of how the methodology was gender responsive, and as the tools are not provided it is not possible to see how qualitative data were captured. However the evaluation participants are disaggregated by stakeholder group and gender. = 1</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) Mixed methods were used and the sample size appeared adequate. = 2</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) The design relied on the usual range of data sources - document review, KIs and observation - which were adequate for inclusion and data quality. Triangulation and validation were evident. = 3</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) The evaluators managed to consult a diverse range and number of stakeholders, including beneficiaries. = 3</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) There is no discussion of ethical principles. = 0</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) There is a well-referenced subsection of Country Context that is dedicated to gender equality and women's empowerment. It briefly addresses challenges faced by girls, women and people with disabilities, as well as national responses = 3</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) There is some reference made to the input of different stakeholder groups but more could be done (for example, by the use of quotes, case examples and disaggregated data) to more clearly highlight different perspectives. = 2</p> <p>2 c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) There is one sentence under Relevance that states no unanticipated effects were found. However, the evaluation did well to present contextual information affecting results, and it appears some results presented were indeed unintended, or unplanned. = 1</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) The findings are very thorough and flow through to GEEW-related lessons and recommendations for future work. = 3</p>
--	---

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totaling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory).

(**) Scoring uses a four point scale (0-3).

0 = Not at all integrated. Applies when none of the elements under a criterion are met.

1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

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	11	69	20	
Overall assessment level of evaluation report		Good		
	Very good very confident to use	Good confident to use	Fair use with caution	Unsatisfactory not confident to use

- (*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'Analysis and findings' has been assessed as 'Good', enter 40 into 'Good' column).
- (b) Assessment level with highest 'total scoring points' determines 'Overall assessment level of evaluation report'. Write corresponding assessment level in cell (e.g. 'Fair').
- (c) Use 'shading' function to give cells corresponding colour.

If the overall assessment is 'Fair', please explain

- How it can be used?

- What aspects to be cautious about?

Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory

Consideration of significant constraints

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:

Yes No

If yes, please explain: