Organizational unit:	UNFPA Evaluation Office

Year of report: 2020

Title of evaluation report: Somalia Country Programme Evaluation - 2018-2020

Overall quality of report:

Very Good

Date of assessment: 27 February 2021

Overall comments: This is a solid evaluation of the Somalia 2018-2020 Country Programme with each thematic component (Sexual and Reproductive Health and Rights, Adolescent and Youth, Gender Equality and Women Empowerment, and Population Dynamics) being individually assessed against the OECD/DAC key criteria of relevance, effectiveness, efficiency and sustainability in addition to assessing coordination, coverage, and connectedness of the programme as a whole. A thorough evaluation process was carried out in spite of covid restrictions. The report is well structured, although a list of annexes could have been shown in the table of contents to make them quickly accessible. The findings are well presented and supported by both qualitative and quantitative data. The methodology section report could be improved by being more explicit about how ethical considerations were addressed in the evaluation process. A further concern is that the main body of the document, at 111 pages, significantly exceeds the maximum length of 70 pages for CPEs, mainly due to an extensive context section and detailed findings. Both the Conclusions and Recommendations sections of the evaluation are well formulated and presented.

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 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 structured in a logical way with a clear distinction between analysis/findings, conclusions and recommendations. There are minimal errors.	
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 111 pages, which is 41 pages beyond the 70 page limit 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 contain the required material.	
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 executive summary is complete as a stand-alone document, but the recommendations could be more clearly presented if numbered and separated into strategic and programmatic recommendation rather than being consolidated into one paragraph.	

5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The length is 6 pages, which goes beyond the maximum length of five pages.	
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	There is a extensive description of the country background in terms of the development, institutional and programmatic context for each of the four programmatic areas as well as overall constraints.	
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The report discusses and assesses the intervention logic / theory of change. The ToC was reconstructed by the evaluation team and appears thorough.	
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 evaluation framework is described in the text but Annex 5 provides a very complete showing of the questions, assumptions, indicators, data sources and methods of data collection.	
4. Are the tools for data collection described and their choice justified?	Yes	The evaluation report identifies the data collection tools and provides the rationale for their selection.	
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	There is a comprehensive stakeholder map by programme in Annex 7 prepared in the early stage of the evaluation. In the section on Methodology the evaluation states "the evaluation was based on an inclusive, transparent and participatory approach" which included "preparation of a second draft CPE report based on review comments of the UNFPA CO, Regional Office and the Evaluation Reference Group.	
6. Are the methods for analysis clearly described for all types of data?	Yes	The report describes the processes for managing and analysing data; these included contribution, content and trend analysis, and the use of descriptive statistics.	
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	Four key methodological limitations and their mitigation strategies are described.	
8. Is the sampling strategy described?	Partial	The evaluation used a purposive sampling strategy, whereby UNFPA stakeholders were selected from all over the country where the programme was implemented according to specified criteria. The evaluators do not include how the universe was determined.	
9. Does the methodology enable the collection and analysis of disaggregated data?	Partial	Although some disaggregated data is presented in the report (including in the country context section), the methodology does not indicate how this was done. In terms of primary data collection, the report text does not include the total number of stakeholders consulted (although the list can be found in the annex), or indicate the percent of women or different stakeholder groups participating. The annexed data collection tools also do not have a field for type or gender of respondent.	

10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	Cross-cutting issues dealing with gender, humanitarian/emergency responses, coordination and connectedness are integrated into the design.
3. Reliability of Data	Yes No Partial	Assessment Level: **Fair**
To ensure quality of data and robust data collection processes		
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation team validated collected data on a regular basis, with quantitative and qualitative data from primary and secondary data being systematically assessed and findings from different sources being triangulated.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	The evaluation was careful to identify and describe the various sources and all were used, both qualitative and quantitative. The report does not fully discuss the reliability (or lack thereof) of both.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The evaluation report notes ethical aspect were in line with the UN Evaluation Group Code of Conduct, Ethical Guidelines and Norms and Standards. Ethical practices mentioned included respecting social distancing, ensuring informant confidentiality, accommodating non-English speakers, and having separate focus group discussions for men and women.
4. Analysis and Findings	Yes No Partial	Assessment Level: **Very good**
To ensure sound analysis and credible findings		
1. Are the findings substantiated by evidence?	Yes	The findings in the evaluation report are supported by sources of evidence in the text, in footnotes and in the annexes.
2. Is the basis for interpretations carefully described?	Yes	In each finding, the basis on which the interpretation was made is shown in detail. Terms commonly used include "interviews with specifically designated stakeholders" and "document reviews revealed … '".
3. Is the analysis presented against the evaluation questions?	Yes	Each of the substantive findings are structured according to the main criteria and start with the evaluation questions and then give the specific findings in summary form before providing the detailed evidence supporting those results.
4. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Where causal connections can be observed, they are reported. The evaluators looked for any unintended outcomes.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings analyze and show differences across relevant outcomes for diverse target groups.
6. Is the analysis presented against contextual factors?	Yes	The findings are presented in the evaluation report within contextual factors at the global, national, regional or local levels as appropriate that affect performance.

7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	In addition to elaborating on cross-cutting issues such as equity and vulnerability, disability inclusion, and gender equality, the evaluation also addresses other types of cross-cutting themes including humanitarian/ emergency responses, coordination and connectedness.
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	Conclusions are based on the analysis of findings and clearly flow from the evidence. Each conclusion is followed by the number of the respective evaluation questions contained in the text of the finding section. They are organized as strategic conclusions and programmatic conclusions.
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 conclusion statements are succinctly stated with supporting text and are effective in conveying the sense of the issues and findings. They also reflect appropriate cross-cutting issues.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias.
6. Recommendations	Yes No Partial	Assessment Level: **Fair**
To ensure the usefulness and clarity of recommendations		
1. Do recommendations flow logically from conclusions?	Yes	Each of the recommendations, at the strategic and programmatic levels, are drawn from the conclusions from which they flowed and they are cross-referenced accordingly.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	The intended users and the operational requirements are shown but not the budgetary 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	They appear balanced and impartial. They also address the need to deliberately target and adequately address the needs of more marginalized and vulnerable populations, including those with disabilities.
4. Are the recommendations prioritized?	Yes	The recommendations are prioritized into high and medium priority.
7. Gender	0 1 2 3 (**)	Assessment Level: **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?	3	**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 specifically mentioned in the objectives but is included as part of the thematic scope of the evaluation process. = 2 **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)** HRGE is mainstreamed under Relevance, Effectiveness and Sustainability = 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)** This is addressed in EQ5. = 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)** This is taken up to some extent in the evaluation matrix = 2
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	2	**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)** It is noted that UNEG and UNFPA guidance on integrating HRGE into the evaluation process was followed and that data was collected with a gender lens. There could be more clarity on the ways this was done, including how disaggregated data was collected. = 2 **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)** A mixed methods and participatory approach is used. It is difficult to ascertain whether the sample size was adequate = 2 **c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)** Sources are diverse as data was collected from an adequate range of stakeholder groups which included focus groups with rights holders. = 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)** Rights holders were included but it is difficult to determine if these included the most vulnerable = 1 **e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3)** There are several examples of how ethical considerations were applied = 3
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	**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 focuses on HRGE and social issues, and includes a significant amount of gender and age-disaggregated data. = 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)** This type of analysis is not presented for rights holders, although perspectives of staff and implementing partners are shown. Overall, the evaluation does well in illuminating ways in which the CP addressed the needs of a range of social groups including those with disabilities = 2 **c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)** The discussion guide includes a question on unanticipated effects but this is not discussed in the report text. = 1 **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)** Multiple recommendations address GEEW. = 3

(*) 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).

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	51	20	29	
Overall assessment level of evaluation report	Very Good			
	Very good	Good	Fair	Unsatisfactory

(*) (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.

OVERALL COMMENTS: Please explain the overall assessment.

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: