

Organizational unit: **Year of report:** 2017

Title of evaluation report: Djibouti 4th country programme evaluation

Overall quality of report: Good **Date of assessment:** 13 September 2017

Overall comments: Overall, the evaluation report is good. The design of the evaluation was good and most of the evaluation limitations were identified (although they were mostly related to data availability). The evaluators tried to reconstruct the program theory of change but failed to link the intervention and expected results to any assumption and/or associated risk. While gender and issues of vulnerability have been integrated in the design, the report does not always provide data disaggregated by gender. The conclusions flow from the findings and the recommendations flow from the conclusions, however the recommendations have not been prioritized and no timeline for their implementation has been provided.

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	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>		
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Fair
<i>To ensure the report is comprehensive and user-friendly</i>			
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?	Yes	The report is easy to read and understand but has several grammatical/vocabulary 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 too long : 153 pages in total and 102 pages excluding the annexes.	
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 report is structured in a logical way. Although there are no lessons learned, there is a clear distinction between the analysis/findings, conclusions and recommendations section.	
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	The annexes are mostly complete. They include the evaluation matrix, the TORs, a bibliography, the list of people interviewed, the data collection tools and the interview guides. There is, however, no information on the stakeholder consultation process.	
<i>Executive summary</i>			
5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?	Yes	The report does include an executive summary as a stand-alone section and does present the main findings of the report.	
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 executive summary has a clear structure. It includes all the main sections of the report.	
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	No	The executive summary is too long (7 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	The report does describe the target audience of the evaluation: the UNFPA Country and Regional Office of the Arab States, donors, development partners, government counterparts and civil society.	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The report gives extensive information on the development and institutional context of the evaluation. The constraints that have been described are mainly related to the deficit of data and the measures that were taken to overcome them.	
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	The evaluators tried to describe the underlying theory of change/intervention logic using a table in the report. That analysis is not complete though as the adequacy of the choices made has not been sufficiently discussed. In addition, the assumptions and risks of the program have not been completely discussed.	
<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 framework is described in the text and in the evaluation matrix. The evaluation matrix does include the evaluation questions, indicators and data sources and methods for data collection but no assumptions has been included.	
5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection have been described. Their choice is justified given the limitations in accessing the information.	
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	All the stakeholders have been identified. The report does talk about the consultation process but nothing indicates that the stakeholders have been consulted regarding the recommendations.	
7. Are the methods for analysis clearly described for all types of data?	Yes	The methods for data analysis have been described for both the qualitative and quantitative data that were gathered.	
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The evaluation's limitations were described (all were related to lack of data availability) and the evaluators tried to identify the impact of this on the evaluation. What was done to overcome them has also been described.	
9. Is the sampling strategy described?	Yes	Yes the sampling strategy has been described in the report.	
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	To the extent that data existed (and this was a limitation for several questions), data could be disaggregated and it was.	
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The design did include questions related to gender, human rights and vulnerability so that these issues could be addressed properly through collection of appropriate data.	

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	Triangulations was done where it was possible. The evaluation relied on several data sources and the information from each of those sources has been used as appropriate. There were few cases where triangulation was difficult.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation used several reliable data sources that were mainly qualitative as stated in the methodology.
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	Data limitations were made clear as were the steps taken to minimize their effect on the analysis.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	There is partial evidence that data has been collected with a sensitivity to issues of discrimination. While the methodology does identify several socially-disadvantaged groups, it has not always been clear how these groups were taken into account in data collection and the information that has been reported does not always specify those groups.

4. Analysis and Findings	Yes No Partial	Assessment Level:	Good
---------------------------------	----------------------	-------------------	-------------

To ensure sound analysis and credible findings

1. Are the findings substantiated by evidence?	Yes	In each case, following the evaluation questions, the evaluators have been careful to show the evidence on which their findings are based.
2. Is the basis for interpretations carefully described?	Yes	The evaluators were meticulous in showing the basis for their interpretations of the data. They sought to respond to each evaluation question.
3. Is the analysis presented against the evaluation questions?	Yes	All the analysis has been presented against the evaluation questions.
4. Is the analysis transparent about the sources and quality of data?	Yes	The evaluators in each case looked for and found causal connections between UNFPA support and observed changes. The evaluators were transparent about the quality of data describe the data limitations and what they did to overcome them.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Partial	Because of the partial description of the intervention logic, it has not always been possible to credibly link the interventions to the end results. The contribution of other development actors in the same domain of intervention has not been determined either.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The different programs were often group specific (vulnerable women, youth, girls, victims of violence, refugees) and in that context the report showed the differences.
7. Is the analysis presented against contextual factors?	Yes	The country context was extensively described and the evaluators did try to link the results to the context.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The contextual issues were built into the design and then into the analysis which addresses gender issues fully, and, to a lesser extent, vulnerability issues. Other than reproductive rights, human rights issues were not addressed.

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	The conclusions are organized by evaluation question and in each case summarize well the findings so that the conclusions flow clearly from them.
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 do not always go beyond the findings and reflect upon broader implications, and, as such an understanding of the underlying issues was not always provided.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusion appear to convey the unbiased judgement of the evaluators. There is no evidence that the evaluators were unfair in making their conclusions.

6. Recommendations	Yes No Partial	Assessment Level:	Fair
---------------------------	----------------------	-------------------	-------------

<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	The recommendations flow logically from the conclusions. There is one recommendation for each conclusion which flows from the nature of the conclusions.
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	While the recommendations are clear in that they define the action to be taken, the intended users/implementers are not identified nor are the timeframe or cost implications.
3. Do recommendations appear balanced and impartial?	Yes	The recommendations appear balanced and impartial.
4. Is a timeframe for implementation proposed?	No	No timeframe has been proposed for their implementation.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	No	The recommendations have not been prioritized.

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?	2	The design includes gender as one of the main foci and GEEW is integrated in the evaluation scope of analysis. However, GEEW data have not always been collected.
2. Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved?	2	The evaluation criteria and evaluation questions specifically address how GEEW has been integrated in the programme implementation and results achieved.
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	The tools that were used and the analysis techniques are appropriate to collect and report gender-disaggregated data. However, though the evaluators collected gender disaggregated data, the analysis does not always reflect on the differentiated outcomes by gender.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	The findings and conclusions reflect a good gender analysis, within the limits of the data.

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

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		69	31	
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

The analysis was designed and was thorough, meriting a good. It was less good in the recommendations which lacked time-horizons and priorities.

Consideration of significant constraints

The quality of this evaluation report has been hampered by exceptionally difficult circumstances: Yes No

If yes, please explain: