

**Organizational unit:** \_\_\_\_\_ **Year of report:** 2018

**Title of evaluation report:** Evaluation indépendante du 7eme programme de cooperation Mali - UNFPA 2015 - 2019

**Overall quality of report:** **Very Good** **Date of assessment:** 11 December 2018

**Overall comments:** Overall, the evaluation report is very good. It is comprehensive and written in an accessible language, with the design and methodology sound and the methods of analysis clearly described. The evaluation a combination of data collection methods – including document review/analysis, interviews and observations – to obtain data that was then used to clearly illustrate which aspects of the programme worked and those that needed to be improved. Notably, in its assessment of effectiveness, the evaluation used a particularly innovative method of analysis, comparing observed results with expected (ideal) results to be able to assess the extent of progress. The main stakeholders were mapped and the limitations of the evaluations were clearly spelled out. While the recommendations flow logically from the conclusions, they have not been costed. The evaluation was gender responsive, including in the tools and methods used.

**Assessment Levels**

- Very Good** (blue box): strong, above average, best practice
- Good** (green box): satisfactory, respectable
- Fair** (yellow box): with some weaknesses, still acceptable
- Unsatisfactory** (red box): 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>		
<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure the report is comprehensive and user-friendly</i>			
<b>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?</b>	Yes	The report is easy to read and written in an accessible language with minimal grammatical, spelling and punctuation errors.	
<b>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)</b>	Yes	The body of the report is 52-page long for a total of 69 pages, excluding the annexes. The report therefore is of a reasonable length.	
<b>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)?</b>	Yes	The report is structured in a logical fashion. There is a clear distinction between analysis/findings, conclusions and recommendations.	
<b>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?</b>	Partial	The annexes are a separate file and include the TORs, the list of interviewees, the evaluation matrix, the data collection tools and bibliography. However, the annexes do not include information on the stakeholder consultation process.	
<i>Executive summary</i>			
<b>5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?</b>	Yes	The executive summary is written as a stand-alone section and presents the main results of the evaluation.	
<b>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)?</b>	Partial	The structure of the executive summary is clear. It includes the Objectives, the methodology, the main conclusions and the recommendations. The intended audience of the evaluation and the description of the interventions however are not included in the executive summary.	
<b>7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	Yes	The executive summary is concise at 5-pages.	

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level:	<b>Good</b>
<i>To ensure that the evaluation is put within its context</i>			
<b>1. Does the evaluation describe the target audience for the evaluation?</b>	Yes	The target audience for the evaluation (the country team and the government counterparts) is clear from the report itself.	
<b>2. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	The development and institutional context of the evaluation is clearly described in the report. The constraints have also been identified. For example, there are several places that the evaluation team was not able to reach because of the bad road network. Insecurity has also prevented the team from visiting many northern communities.	
<b>3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?</b>	Yes	The intervention logic is shown on p. 22 of the evaluation report in considerable detail. In the analysis/description that precedes it, the context, the main health-related development issues and the response from UNFPA have been described.	
<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 clearly described in the text . The evaluation matrix establishes the evaluation questions, assumptions, indicators, data sources and methods of data collection.
5. Are the tools for data collection described and their choice justified?	Yes	The evaluation uses document analysis, group and individual interviews (together with observation) as the main data collections methods. The choice of the tools have been justified. They were appropriately customized and adapted according to the evaluation question and the interviewees/observations.
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	The stakeholders have been identified. While an evaluation reference group chaired by the ministry of foreign affairs was put in place, the overall stakeholder consultation process has only been partially described.
7. Are the methods for analysis clearly described for all types of data?	Yes	The methods for analysis have been described for all types of data. For both quantitative and qualitative data the consultants have opted to calculate the theoretical achievement rate that they compared against the actual achievement rate. Data from individual interviews and group discussions have been triangulated as shown on page 8 and 9.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Partial	The methodological limitations have been acknowledged and their effect on the evaluation has also been identified. However, the evaluation only partially indicates how they were overcome. For example, several constraints/limitations were listed for which a mitigating action/solution was not described (i.e. some people to be interviewed were not available due to the electoral process, some places could not be visited, students were on holidays).
9. Is the sampling strategy described?	Yes	The interviews and visits are based on a purposive sample that is itself based on a clear stratification of factors.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	For example, the evaluation disaggregated interviews by gender, type of beneficiary or implementer, region, type of institution which allowed a reasonable use of disaggregated data.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The design and methodology are appropriate for assessing cross cutting issues such as human rights, gender equality and vulnerability. All groups of beneficiaries have been interviewed across all the program areas. A rapid gender analysis has been done to understand the issues faced by both women and youth; and the methodology has been adapted to capture the views of each group.

<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To ensure quality of data and robust data collection processes</i>			
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation triangulated the data that was collected in an appropriate fashion.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation identified and made use of reliable qualitative and quantitative data sources	
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	All the evaluation limitations have been identified but the report only indicates how some of them have been addressed. For example while several places had security issues, the evaluators indicated that they used security convoys to reach some of them and did not indicate how they interviewed people in places they were not able to reach. The absence of some political leaders were noted but no indication has been given regarding how the evaluators gathered their opinions.	
4. 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 have been collected with a sensitivity to issues such as discrimination and other ethical considerations. For example, in the description of the technical preparation for field data collection, one approach included "La mise en confiance des interlocuteurs par une mise en évidence du caractère indépendant de l'évaluation, l'assurance de la confidentialité des informations recueillies et la demande de leur consentement éclairé."	

<b>4. Analysis and Findings</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	The findings have been substantiated by evidence. For example, evidence - such as quotes from interviews and official statistics - is triangulated and ground the findings.	
2. Is the basis for interpretations carefully described?	Yes	The presentation of findings was highly detailed so that the basis for interpretation was clear in all cases.	
3. Is the analysis presented against the evaluation questions?	Yes	The analysis has been presented against the evaluation questions. All the evaluation criteria have been addressed.	
4. Is the analysis transparent about the sources and quality of data?	Yes	The analysis is transparent about the sources and quality of data.	

5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Throughout the findings, the evaluators were careful to show the connections between UNFPA's interventions and the results obtained. Particularly in the findings on effectiveness, the evaluators showed the clear connection between interventions and what were called outputs (which would be outcomes if the UNFPA interventions were considered the programme's output). There were almost no unintended outcomes noted.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis shows different outcomes for different target groups. Youth and women have been specifically singled out. Table 4.1 pages 26-29 details how the different programs target different groups. To highlight the flexibility and responsiveness of the program, the evaluators have shown how the activities for youth and midwives have been adapted to respond to the ebola crisis, and the analysis under achievements of indicators 1 and 2 show the results of the program for youth and vulnerable women such as sex workers.
7. Is the analysis presented against contextual factors?	Yes	Yes, the context in which the analysis/findings are embedded is described, including, importantly, the conflict.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The analysis includes information on how cross cutting issues such as vulnerability, gender equality and human rights have been integrated. There are specific indicators for youth and vulnerable groups such as women and sex workers and the evaluators have assessed the program performance with regard to each group. SGBV is also very common in Mali and the analysis has shown the progress made in the fight against it.

<b>5. Conclusions</b>	Yes No Partial	Assessment Level:	<b>Good</b>
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	The 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?	Partial	The conclusions partially go beyond the findings. They convey a reasonable understanding of the underlying issues of the programme in Mali. The conclusions are short and do not always provide much detail about the issues that have been analyzed. For example the evaluators could only say the program adequately contributed to the UNDAF objectives without being more specific when asked to what extent the program contributes to the UNDAF planned objectives.	
3. Do the conclusions appear to convey the evaluators' unbiased judgment?	Yes	The conclusions appear to convey the evaluators unbiased judgment.	

<b>6. Recommendations</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The recommendations flow directly from 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	The recommendations are clear and priority levels given; however, the intended users and financial implications were not identified.	
3. Do recommendations appear balanced and impartial?	Yes	The recommendations appear balanced and impartial.	
4. Is a timeframe for implementation proposed?	Yes	They are either to be implemented in the last year of the current programme or in the next (8th) programme.	
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	They are given either high or medium priority and are specific enough to be followed up.	

<b>7. Gender</b>	0 1 2 3	Assessment Level:	<b>Good</b>
<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?	3	GEEW was integrated in the evaluation scope of analysis and the indicators allowed for GEEW-related data to be 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 questions partially address how GEEW has been integrated into the planning and implementation of the interventions. For example, evaluation question 2 specifically refers to the effectiveness of the program and the results achieved vis a vis both men and women (binary gender disaggregation).	
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	3	A gender-responsive evaluation methodology, methods and techniques have been used. For example, in data collections methods, in most cases, men and women were interviewed separately. The evaluators successfully identified the issues that affect men, women and youth separately and in each case they have shown the different results for each group.	

4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	Although the findings included a rapid gender-analysis about, inter alia, women, youth, sex workers, and other vulnerable groups, the conclusions and recommendations do not make specific references to youth, sex workers and other vulnerable groups, rendering the gender analysis partial.
---	---	---

(\*) 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		
<b>Total scoring points</b>	<b>51</b>	<b>31</b>	<b>18</b>	
<b>Overall assessment level of evaluation report</b>	<b>Very Good</b>			
	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?  
The report is very clear on the extent to which targets were achieved, although these were mostly output targets. Its fair rating had to do with the weighting in several criteria.

• What aspects to be cautious about?

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

The conduct of the evaluation was extremely competent in design and execution and for that reason merited a very good rating.

**Consideration of significant constraints**

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

If yes, please explain: