

**Organizational unit:**

**Year of report:** 2016

**Title of evaluation report:** **EVALUATION FINALE DU 7<sup>ème</sup> PROGRAMME DE COOPERATION BURUNDI – UNFPA 2010 – 2015**

**Overall quality of report:**

**Good**

**Date of assessment:** 9 December 2016

**Overall comments:**

The evaluation covers the country programme. While its presentation is stylistically different and there are some weaknesses in how the methodology was designed and implemented, the findings are solid and are based on clearly expressed evidence, the conclusions flow from this and the recommendations to the country office for the next planning period are clear.

#### 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

<b>Quality Assessment Criteria</b>	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>	
<b>I. Structure and Clarity of Reporting</b>  <i>To ensure the report is comprehensive and user-friendly</i> <ul style="list-style-type: none"> <li>• Is the report easy to read and understand (i.e. written in an accessible non-technical language appropriate for the intended audience)?</li> <li>• Is the report focused and to the point (e.g. not too lengthy)?</li> <li>• Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?</li> <li>• Do the annexes contain – at a minimum – the ToRs; a bibliography, a list of interviewees, the evaluation matrix and methodological tools used (e.g. interview guides; focus group notes, outline of surveys)?</li> </ul> <i>Executive summary</i> <ul style="list-style-type: none"> <li>• Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?</li> <li>• 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)?</li> <li>• Is the executive summary reasonably concise (e.g. with a maximum length of 5-10 pages)?</li> </ul>	<b>Assessment Level:</b>	<b>Fair</b>
	<b>Comment:</b> The evaluation has the expected content. The Executive Summary is the correct length, but is not a self-standing text since it is structured almost like an outline, where all of the requested sections are present but are not in an easy-to-read format.	

<p><b>2. Design and Methodology</b></p>	<p>Assessment Level: <b>Fair</b></p>
<p><i>To ensure that the evaluation is put within its context</i></p> <ul style="list-style-type: none"> <li>• Does the evaluation describe whether the evaluation is for accountability and/or learning purposes?</li> <li>• Does the evaluation describe the target audience for the evaluation?</li> <li>• Is the development and institutional context of the evaluation clearly described?</li> <li>• Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change?</li> <li>• Does the evaluation explain any constraints and/or general limitations?</li> </ul> <p><i>To ensure a rigorous design and methodology</i></p> <ul style="list-style-type: none"> <li>• Is the evaluation approach and framework clearly described? Does it establish the evaluation questions, assumptions, indicators, data sources and methods for data collection?</li> <li>• Were the methods chosen appropriate for addressing the evaluation questions? Are the tools for data collection described and justified?</li> <li>• Is the methods for analysis clearly described?</li> <li>• Are methodological limitations acknowledged and their impact on the evaluation described? (Does it discuss how any bias has been overcome?)</li> <li>• Is the sampling strategy described? Does the design include validation techniques?</li> <li>• Is there evidence of involvement of stakeholders in the evaluation design? (Is there a comprehensive/credible stakeholder map?)</li> <li>• Does the methodology enable the collection and analysis of disaggregated data?</li> <li>• Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?</li> </ul>	<p>Comment: The evaluation's purpose and context are well-expressed, and the target is probably the country office. In different parts, it reconstructs the intervention logic and it indicates constraints. However, it does not clearly describe the data collection process, other than that it would be based on the menu of documents, interviews and focus groups. The persons who were interviewed or are in the focus groups are listed in the annex, but are not shown in the text and the basis for selection (other than that they are in one or another office, like UNFPA) is not described.</p>

<b>3. Reliability of Data</b>	Assessment Level:	<b>Fair</b>
<p><i>To ensure quality of data and robust data collection processes</i></p> <ul style="list-style-type: none"> <li>• Did the evaluation triangulate all data collected?</li> <li>• Did the evaluation clearly identify and make use of qualitative and quantitative data sources?</li> <li>• Did the evaluation make explicit any possible issues (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues? I.e. did the evaluation make explicit possible limitations of the data collected?</li> <li>• Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</li> <li>• Is there adequate gender disaggregation of data? And if this has not been possible, is it explained?</li> <li>• Does the evaluation make explicit the level of involvement of different stakeholders in the different phases of the evaluation process?</li> </ul>	<p>Comment: The evaluators state that they triangulated the data and there is indication of this. They used both quantitative and qualitative data. Little is said about limitations with regard to the data collected. Apart from a reference to the UNEG ethical norms and standards, the report does not provide clear evidence that data has been collected with sensitivity to issues of discrimination and other ethical considerations.</p>	

<b>4. Analysis and Findings</b>	Assessment Level:	<b>Good</b>
<p><i>To ensure sound analysis</i></p> <ul style="list-style-type: none"> <li>• Is information analysed and interpreted systematically and logically?</li> <li>• Are the interpretations based on carefully described assumptions?</li> <li>• Is the analysis presented against the evaluation questions?</li> <li>• Is the analysis transparent about the sources and quality of data?</li> <li>• Are possible cause and effect links between an intervention and its end results explained?</li> <li>• Where possible, is the analysis disaggregated to show different outcomes between different target groups?</li> <li>• Are unintended results identified?</li> <li>• Is the analysis presented against contextual factors?</li> <li>• Does the analysis include reflection of the views of different stakeholders (reflecting diverse interests)? E.g. how were possible divergent opinions treated in the analysis?</li> <li>• Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?</li> </ul> <p><i>To ensure credible findings</i></p> <ul style="list-style-type: none"> <li>• Can evidence be traced through the analysis into findings? E.g. are the findings substantiated by evidence?</li> <li>• Do findings follow logically from the analysis?</li> <li>• Is the analysis of cross-cutting issues integrated in the findings?</li> </ul>	<p>Comment: The information is presented by evaluation question and generally indicates findings supported by data. Assumptions underlying the interpretations are clearly described, and contextual factors well identified. In a number of cases related to effectiveness, the connection between observed results and UNFPA contribution is not obvious. For example, the evaluators report that : « <i>Le nombre de structures sanitaires assurant les fonctions SONUB est passé de 5 en 2010 à 19 en 2015 au niveau national. La disponibilité de l'offre par fonction SONUC était de 17 en 2010 et de 46 en 2015 (Enquête SPSR de 2015)</i> » but the contribution from UNFPA to this result is not analyzed. In other sections, however, the causal connections are clear and the findings well-supported and balanced, by showing that UNFPA-supported training, for example, improved participation of women in political processes. Gender generally is a major issue and is well-addressed in the analysis and findings. The findings reflect different perspectives of stakeholders in the presentation.</p>	

<b>5. Conclusions</b>	Assessment Level:	<b>Good</b>
<p><i>To assess the validity of conclusions</i></p> <ul style="list-style-type: none"> <li>• Are conclusions credible and clearly related to the findings?</li> <li>• Are the conclusions demonstrating an appropriate level of analytical abstraction?</li> <li>• Are conclusions conveying the evaluators' unbiased judgement of the intervention?</li> </ul>	<p>Comment: Although it is drafted in an outline style which makes its reading somewhat difficult, the conclusions flow directly from the findings and show clearly what the evaluation has found out about the programme. There is a very effective SWOT table that covers the main areas and shows clearly what has worked and what has been less effective.</p>	

<b>6. Recommendations</b>	Assessment Level:	<b>Good</b>
<p><i>To ensure the usefulness and clarity of recommendations</i></p> <ul style="list-style-type: none"> <li>• Do recommendations flow logically from conclusions?</li> <li>• Are the recommendations sufficiently clear, targeted at the intended users and operationally-feasible?</li> <li>• Do recommendations reflect stakeholders' consultations whilst remaining balanced and impartial?</li> <li>• Is the number of recommendations manageable?</li> <li>• Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?</li> </ul>	<p>Comment: Again, the formatting is unusual (completely tabular in this case), but the recommendations are shown clearly to flow from the conclusions and are assigned a priority. All are addressed to the country office and are intended to affect the next country programme.</p>	

<b>7. Gender</b>	Assessment Level:	<b>Fair</b>
<p><i>To assess the integration of Gender Equality and Empowerment of Women (GEEW)<sup>1</sup></i></p> <ul style="list-style-type: none"> <li>• Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?</li> <li>• Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved?</li> <li>• Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?</li> <li>• Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</li> </ul>	<p>Comment: The programme gives a high importance to gender and it is built carefully into UNFPA work.</p> <p>Apart from one evaluation question (EQ2, on effectiveness) which includes an assumption relating to gender, GEEW is not integrated in the evaluation criteria and questions. Indicators selected are not conducive to and adequate assessment of the extent to which GEEW has been integrated into the design, planning, and the implementation of the interventions. Findings, conclusions and recommendations do address gender issues, however they do not reflect a proper gender analysis.</p>	

---

<sup>1</sup> 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). One question is if this criteria should be included in the overall evaluation quality assessment grid, or form a separate column and be assessed on its own.

### 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	
<b>Total scoring points</b>		<b>62</b>	<b>38</b>	
<b>Overall assessment level of evaluation report</b>		<b>Good</b>		
	<b>Very good</b> ➔ very confident to use	<b>Good</b> ➔ confident to use	<b>Fair</b> ➔ use with caution	<b>Unsatisfactory</b> ➔ not confident to use

(\*) **(a)** Insert scoring points associated with criteria in corresponding column (e.g. - if 'finding and analysis' 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<sup>2</sup>:**

- How it can be used?
- What aspects to be cautious about?

**Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory<sup>3</sup>:**

**Consideration of significant constraints<sup>4</sup>**

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

yes

no

If yes, please explain:

---

<sup>2</sup> The purpose here is to clarify in what way the report can be used. This in order to assist the elaboration of a relevant Management Response and the wider use of the evaluation findings back into programming. When a report has been assessed as Fair, it is obligatory to fill this text box in.

<sup>3</sup> The purpose is, where relevant, to clarify for example severe unbalances in the report (for example, the report is good overall but recommendations very weak). Is optional to fill in.

<sup>4</sup> E.g. this should only be used in case of significant events that has severely hampering the evaluation process like natural disasters, evaluators falling sick, unexpected significant travel restrictions, etc. More 'normal' limitations should be mentioned under relevant section above.