

# EQA for UNFPA Zimbabwe 6<sup>th</sup> Country Programme

## Evaluation (2012-2015)



**Title of Evaluation Report: End of the Government of Zimbabwe & UNFPA 6th Country Programme (2012-2015) Evaluation**

**Overall Quality Rating: Good**

**Overall Assessment:** The evaluation report is presented in a user-friendly manner and is well written. In its design and methodology, the evaluation is careful to define results so that the contribution of UNFPA is made visible. Its findings show the extent to which UNFPA supported activities have contributed to the achievement of country programme outputs. Findings are based on evidence, including that gathered by the evaluators own data collection techniques. The findings, conclusions, and recommendations are well organized. The conclusions are consistent with the findings, and the evaluation team has been careful and cautious not to extend their conclusions beyond what the findings permit and by so doing clearly convey an unbiased judgment of the intervention. The recommendations are strategic, targeted and operational.

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p><b>I. Structure and Clarity of Reporting</b>  <i>To ensure report is user-friendly, comprehensive, logically structured and drafted in accordance with international standards.</i>                      Checklist of minimum content and sequence required for structure:</p> <ul style="list-style-type: none"> <li>• i) Acronyms; ii) Exec Summary; iii) Introduction; iv) Methodology including Approach and Limitations; v) Context; vi) Findings/Analysis; vii) Conclusions; viii) Recommendations; ix) Transferable Lessons Learned (where applicable)</li> <li>• Minimum requirements for Annexes: ToRs; Bibliography; List of interviewees; Methodological instruments used.</li> </ul>	<p><b>Good</b>                      The report is well written and user friendly, comprehensive, and logically structured. The report contains all the items listed in the minimum checklist in both the critical elements of the text and the annexes.                      The Introduction includes an overview of the evaluation purpose, scope and approach, the evaluation criteria and key questions, although the placement of the lessons learned and good practices (a major objective of the evaluation) is not easily identified. Context includes two separate sections – one on the country context and the other on the UN/UNFPA response/strategy that provide a solid basis for the evaluation. A map of the country showing regional jurisdiction and a Key Facts Table on Zimbabwe provide useful additional context. The Annexes are complete.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p><b>2. Executive Summary</b>  <i>To provide an overview of the evaluation, written as a stand-alone section and presenting main results of the evaluation.</i>            Structure (paragraph equates to half page max):</p> <ul style="list-style-type: none"> <li>i) Purpose, including intended audience(s); ii) Objectives and Brief description of intervention (1 para); iii) Methodology (1 para); iv) Main Conclusions (1 para); v) Recommendations (1 para). Maximum length 3-4 page.</li> </ul>	<p><b>Good</b>            The report contains a well-written summary that presents the main results and serves as a stand-alone section. This section stays within the prescribed page length (4 pages), and all of the basic elements are included. The main conclusions and recommendations presented are clear and well-written, however they are longer than the recommended length for those two subsections, and the prioritization of recommendations is not immediately clear.</p>			
<p><b>3. Design and Methodology</b>  <i>To provide a clear explanation of the following elements/tools</i>            Minimum content and sequence:</p> <ul style="list-style-type: none"> <li>Explanation of methodological choice, including constraints and limitations;</li> <li>Techniques and Tools for data collection provided in a detailed manner;</li> <li>Triangulation systematically applied throughout the evaluation;</li> <li>Details of participatory stakeholders' consultation process are provided;</li> <li>Details on how cross-cutting issues (vulnerable groups, youth, gender, equality) were addressed in the design and the conduct of the evaluation.</li> </ul>	<p><b>Good</b>            The methodological approach was well described along with the techniques and tools for data collection, and details of the participatory stakeholders' consultation process were included and sufficient. The use of Indicator Matrix and Evaluation Matrix in the Annexes includes the team's caveat that "RFF indicators and results chain logic are insufficient to allow the desired analysis of intended results from the intervention" (p 4).            Some of the elements that were not well described included cross-cutting issues. The constraints/limitations provided in the report primarily focus on resources (the change in team leader early in the evaluation and the consequent shortage of time for the evaluation) but do not raise issues derived from of the methodological choice. A tool for site visits, an important component for recording field observations/data in assessing programme effectiveness, was absent, although as the team noted (p 4) "The Design Report elaborates on each method" i.e., document review, interviews, focus group discussion, and field observations." Also there were no details in the methodology section on how the cross-cutting issues were to be addressed, even though findings about such issues were presented in the Executive Summary and Findings chapters. However, these do not significantly affect the quality of this section.            The sampling methods was a purposive sample and it is stated that "The CO</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	provided the partner list from which to develop stakeholder mapping and draw a sampling frame” and “the respective programme officers advised on stakeholder selection.” Although this raises some concerns about objectivity and independence of the evaluation, the team’s explanation (p 4) of how they balanced the various considerations in developing the sampling provides reassurance.			
<p><b>4. Reliability of Data</b></p> <p><i>To clarify data collection processes and data quality</i></p> <ul style="list-style-type: none"> <li>• Sources of qualitative and quantitative data have been identified;</li> <li>• Credibility of primary (e.g. interviews and focus groups) and secondary (e.g. reports) data established and limitations made explicit;</li> <li>• Disaggregated data by gender has been utilized where necessary.</li> </ul>	<p><b>Good</b></p> <p>The sources of data are identified, albeit often to a limited extent, and the credibility of the data is addressed and limitations made explicit throughout the report. Disaggregated data by gender is used primarily in Context section, and at times in the discussion of findings. The data was carefully derived from documents and supplemented by a well-structured set of interviews that permitted triangulation.</p> <p>The evaluation team highlighted that there is a lack of adherence to RBM in the country programme that led to minimal results data being available for use by the evaluation team and the consequent weakness of outcome data to deal with the effectiveness standard. The team took this into account in framing their findings and conclusions.</p> <p>Regarding use of the Evaluation Matrix, it is included in Annexes but, as noted, the evaluation team indicates it is not supported with results data on UNDAF outcomes, a concern the evaluation team raises early as well as a number of times throughout the report. The evaluation team state: “Both quantitative and qualitative data were captured, the former mainly from databases tracking outputs and experiences” (p.4). The absence of outcome data is duly noted and was taken into account in the evaluators’ findings. An example of the same point elsewhere, the report states “it is a challenge to measure effectiveness of outputs against outcomes, let alone impacts ... and the output indicators in the CPAP are essentially process measures and it is not clear that the theory of change logic is sufficiently robust.” (p 32).</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	Where it is appropriate, the evaluators note “a clear logic between the outcome and outputs and planned activities making for a more clear analysis in this programme, even though they noted “for the remainder of the 6 <sup>th</sup> CP into the 7 <sup>th</sup> , there is a need to collect both quantitative and qualitative data” (p 45).			
<p><b>5. Findings and Analysis</b>  <i>To ensure sound analysis and credible findings</i></p> <p><u>Findings</u></p> <ul style="list-style-type: none"> <li>• Findings stem from rigorous data analysis;</li> <li>• Findings are substantiated by evidence;</li> <li>• Findings are presented in a clear manner</li> </ul> <p><u>Analysis</u></p> <ul style="list-style-type: none"> <li>• Interpretations are based on carefully described assumptions;</li> <li>• Contextual factors are identified.</li> <li>• Cause and effect links between an intervention and its end results (including unintended results) are explained.</li> </ul>	<p><b>Good</b></p> <p>Within the limits of the data that the evaluators noted, the findings flow from the evidence provided. The sources of data for the finding are noted and the findings are presented clearly. Contextual factors and underlying assumptions are often presented in a clear manner: for example, the authors explain how unexpected increases in workload may have affected the provision of programming. The degree to which results were achieved relative to set targets is explained where appropriate; in cases where results were not commensurate with targets, explanation is given.</p> <p>Cause and effect links between the intervention and its end results are carefully mapped, taking into account that what are termed outputs in the country programme and UNDAF are, for UNFPA, outcomes. The evaluation team has noted difficulties they experienced identifying cause and effect links. For example, the evaluation noted that an expected output (training courses) had been done on schedule, but that there was little data on whether the results of the training (improved services that would change behavior of persons using the services) had occurred.</p> <p>In certain cases findings are listed/presented as recommendations. For example, the report specifies that the RRF in the next CP should have stronger results logic and SMART indicators that are more robust, rather than simply using processes and number counts. However, this does not adversely affect the overall quality of the section.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p><b>6. Conclusions</b>  <i>To assess the validity of conclusions</i></p> <ul style="list-style-type: none"> <li>• Conclusions are based on credible findings;</li> <li>• Conclusions are organized in priority order;</li> <li>• Conclusions must convey evaluators' unbiased judgment of the intervention.</li> </ul>	<p><b>Good</b></p> <p>The CPE conveys the evaluation team's unbiased judgment of the intervention, with the report making clear the limitations of the evaluation: for example, the report identifies that "weak and mainly process indicators have limited the CPE in tracking contributions to outcomes and impacts to demonstrate programme effectiveness" (p60). The evaluators are careful not to draw conclusions that cannot be substantiated, making such statements as "it has been difficult to assess the efficiency of resources ... as highlighted above" (p 60), "contributions to the outcomes require a stronger RRF and results chain logic (as noted earlier)." On the other hand, where supported by the findings, they draw conclusions that the "Gender Equality Programme has been the mainstay of efforts to end GBV in Zimbabwe" (p 63). The conclusions were consistent with the findings. The evaluation team has been careful and cautious not to extend their conclusions beyond what the findings permit and by so doing clearly convey an unbiased judgment of the intervention.</p> <p>It is not clear that transferable lessons have been presented; this is not insignificant, as 'transferable lessons' represent the third overall objective of the evaluation. However, overall this is a good quality section.</p>			
<p><b>7. Recommendations</b>  <i>To assess the usefulness and clarity of recommendations</i></p> <ul style="list-style-type: none"> <li>• Recommendations flow logically from conclusions;</li> <li>• Recommendations must be strategic, targeted and operationally-feasible;</li> <li>• Recommendations must take into account stakeholders' consultations whilst remaining impartial;</li> <li>• Recommendations should be presented in priority order</li> </ul>	<p><b>Good</b></p> <p>The recommendations are strategic, targeted and operational, and are organized in a manner consistent with the conclusions chapter.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p><b>8. Meeting Needs</b></p> <p>To ensure that Evaluation Report responds to requirements (scope &amp; evaluation questions/issues/DAC criteria) stated in the ToR (ToR must be annexed to the report). In the event that the ToR do not conform with commonly agreed quality standards, assess if evaluators have highlighted the deficiencies with the ToR.</p>	<p><b>Good</b></p> <p>Building on a clear ToR, the report responds to the requirements in the ToR. However, the report did not identify the lessons learned and good practices of the programme in a way that are easy to find.</p> <p>The evaluators did highlight the “deficiencies” in the ToR particularly that “sufficient operations research is not in place to measure many intended results” (p 4), thereby negating the main purpose of the evaluation “to assess the progress of the 6<sup>th</sup> CP toward achieving the CPAP outputs and the outcomes” (p. 72), particularly the latter. The team reiterated this major constraint at different points in the report.</p>			

Quality assessment criteria (and Multiplying factor *)	Assessment Levels (*)			
	Very good	Good	Poor	Unsatisfactory
1. Structure and clarity of reporting (2)		2		
2. Executive summary (2)		2		
3. Design and methodology (5)		5		
4. Reliability of data (5)		5		
5. Findings and analysis (50)		50		
6. Conclusions (12)		12		
7. Recommendations (12)		12		
8. Meeting needs (12)		12		
<b>TOTAL</b>		100		

(\*) Insert the multiplying factor associated with the criteria in the corresponding column e.g. - if “Finding and Analysis” has been assessed as “good”, please enter the number 50 into the “Good” column. The Assessment level scoring the higher number of points will determine the overall quality of the Report

**OVERALL QUALITY OF REPORT: Good**