

EQA for UNFPA Colombia Country Programme

Evaluation (2008-2014)



Title of Evaluation Report: EVALUACIÓN FINAL DEL V PROGRAMA DE PAÍS (2008 – 2014) FONDO DE POBLACIÓN DE LAS NACIONES UNIDAS EN COLOMBIA

Overall Quality Rating: Good

Overall Assessment: The evaluation examines a complex programme over a six-year period, taking into account a complex context and some changes in orientation during the period. It builds on a thorough terms of reference and determines clearly expected results. Its design and methodology, primarily based on document analysis, but also on extensive interviews from a purposive sample, produces data that can be triangulated to produce credible findings. In this, the evaluation does an excellent job of showing the causal connection between what UNFPA provides and the outputs achieved (in terms of the UNDAF and the UNFPA Country Programme). The volume of findings, in which there is a high level of detail, somewhat loses the larger picture in the many details, however. The conclusions are clear and the recommendations are designed to be clear and implementable, although some terms used lack precision.

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p>I. Structure and Clarity of Reporting <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"> • 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) • Minimum requirements for Annexes: ToRs; Bibliography; List of interviewees; Methodological instruments used. 	<p>Good The report contains all of the necessary sections including the required annexes. However, the report's length does detract slightly from the overall quality.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p>2. Executive Summary <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"> 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. 	<p>Poor The summary is a stand-alone description of the evaluation and is short (two pages). Its main focus is on conclusions and recommendations. The paragraph on methodology, however, does not describe the methodology followed (merely that the methodology was consistent with norms). Furthermore, the summary does not provide a brief description of the intervention, referring only to the three programme areas of the country programme but not listing what the programme areas are.</p>			
<p>3. Design and Methodology <i>To provide a clear explanation of the following elements/tools</i> Minimum content and sequence:</p> <ul style="list-style-type: none"> Explanation of methodological choice, including constraints and limitations; Techniques and Tools for data collection provided in a detailed manner; Triangulation systematically applied throughout the evaluation; Details of participatory stakeholders' consultation process are provided; Details on how cross-cutting issues (vulnerable groups, youth, gender, equality) were addressed in the design and the conduct of the evaluation. 	<p>Good The methodology is well-explained and, on the whole, meets quality requirements. Triangulation, particularly between document analysis and interviews is used consistently throughout. There is clear evidence of stakeholder consultation, specifically the evaluation reference group (GRE), in the establishment of the final evaluation structure as well as in the analysis. Cross-cutting issues are dealt with as major themes and were addressed in the conduct of the evaluation. For example, two semi-structured group interviews were conducted with youth beneficiaries.</p> <p>Limitations in the methodology (such as the need to replace focus groups with semi-structured interviews) are clearly expressed. A large amount of data has been acquired through interviews. A weakness in the design is that the selection of persons to be interviewed was a purposive sample. Interviewees were drawn from the main categories of persons, but those at departmental level were selected by departmental staff based on a criterion of "knowledge of the programme". While this might give good information, it is not clear whether this would be representative of the groups.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p>4. Reliability of Data <i>To clarify data collection processes and data quality</i></p> <ul style="list-style-type: none"> • Sources of qualitative and quantitative data have been identified; • Credibility of primary (e.g. interviews and focus groups) and secondary (e.g. reports) data established and limitations made explicit; • Disaggregated data by gender has been utilized where necessary. 	<p>Good</p> <p>Within the limits of the design, the data used are reliable. Sources are clearly indicated in the text, but are also shown in the annexes. Credibility is clear and the limitations have been described. Efforts were also made to correct for the limitations identified. For example, the report identifies that the indicators for Output 8 were not well-articulated or well-tracked. However, this was taken account of (to some extent) by taking the alignment between the CPAP with the UNFPA Strategic Plan results framework as a reference point, and determining lessons learned from initial implementation to improve the indicators. Gender-disaggregated data have been used. For example, data on interview participants was disaggregated by gender, with 74% identifying as women and 26% as men. Data were also disaggregated by age, ethnicity, geographic area and income level, when available and relevant.</p>			
<p>5. Findings and Analysis <i>To ensure sound analysis and credible findings</i></p> <p><u>Findings</u></p> <ul style="list-style-type: none"> • Findings stem from rigorous data analysis; • Findings are substantiated by evidence; • Findings are presented in a clear manner <p><u>Analysis</u></p> <ul style="list-style-type: none"> • Interpretations are based on carefully described assumptions; • Contextual factors are identified. • Cause and effect links between an intervention and its end results (including unintended results) are explained. 	<p>Good</p> <p>The findings are complete and thorough. They take up the largest part of the report and are structured according to the evaluation questions, within a broader context of strategy, specific programmes, and issues like coordination, monitoring and evaluation. In each, the evaluators have been careful to find causal connections between what support UNFPA has provided and what has resulted. For example, the report identifies the progress made by the country programme in updating, applying, and disseminating standards; similarly, the progress made training Regional Directorates of Health (DTS) is discussed with reference to results (78.5% of municipalities meaningfully integrated components of Sexual and Reproductive Health into their Regional Public Health Plans).</p> <p>The findings respond to the twelve questions that have been posed,</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	<p>including using three questions for effectiveness rather than one. For each, the focus has been on the achievement of what are called output (productos) in the UNDAF and the Country Programme which, for UNFPA would actually be outcomes. This is done very clearly. The evaluators are careful to show how their findings flow from the data.</p> <p>A slight weakness is to note the overall evidence from the interviews rather than breaking it down by type of interviewee, although in most cases this would not be a particular problem since the analysis is done by level (central, departmental) and by subject. The volume of findings reflected in the 54 pages of programmatic analysis, strategic positioning and transversal aspects detracts from the findings by being too detailed and output focused rather than drawing larger conclusions. In fact, the conclusions drawn from these, as noted in the following section, show that the analysis could have had a larger focus.</p>			
<p>6. Conclusions <i>To assess the validity of conclusions</i></p> <ul style="list-style-type: none"> • Conclusions are based on credible findings; • Conclusions are organized in priority order; • Conclusions must convey evaluators' unbiased judgment of the intervention. 	<p>Good</p> <p>The conclusions are organized by number and draw on the findings, but are more condensed. They are clearly based on the findings and there is no evidence of bias on the part of the evaluators.</p>			
<p>7. Recommendations <i>To assess the usefulness and clarity of recommendations</i></p> <ul style="list-style-type: none"> • Recommendations flow logically from conclusions; • Recommendations must be strategic, targeted and operationally-feasible; • Recommendations must take into account stakeholders' consultations whilst remaining impartial; • Recommendations should be presented in priority order 	<p>Good</p> <p>The recommendations flow from the conclusions and from the findings. They are organized by subject area, have been assigned priority and who should implement them and how, showing their operational feasibility. A weakness is that occasionally the recommendations use terms (like strengthen and enhance) whose meaning in practical terms is somewhat ambiguous, therefore making them less actionable.</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p>8. Meeting Needs To ensure that Evaluation Report responds to requirements (scope & 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>Good The report corresponds directly with the ToR, which was very detailed. The evaluators increased the number of questions under effectiveness to consider a more nuanced approach.</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		
TOTAL		98	2	

(*) 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