

Organizational unit: **Armenia Country Office** Year of report: **2019**

Title of evaluation report: **Third UNFPA Country Programme: Armenia 2016-2020 Final Evaluation Report**

Overall quality of report: **Very Good** Date of assessment: \_\_\_\_\_

Overall comments: This is a comprehensive evaluation of the Country Programme with each programme component being individually assessed for relevance, effectiveness, efficiency and sustainability. The findings are well supported by both qualitative and quantitative data, and include illustrative quotes and the perspectives of various stakeholder groups. The report is well structured and is exemplary for its presentation of best practices. These are highlighted in text boxes and appear instructive for the next country programme as well as other initiatives. The report could be improved by being more explicit about how the evaluation process was gender responsive and how ethical considerations were addressed. A further concern is that the main body of the document, at 100 pages, significantly exceeds the maximum length, mainly due to an extensive context section and detailed findings.

<b>Assessment Levels</b>	<b>Very Good</b> strong, above average, best practice	<b>Good</b> satisfactory, respectable	<b>Fair</b> with some weaknesses, still acceptable	<b>Unsatisfactory</b> weak, does not meet minimal quality standards
--------------------------	---	---------------------------------------	--	---

Quality Assessment Criteria	Insert <i>assessment level</i> followed by main <i>comments</i> . (use 'shading' function to give cells corresponding colour)		
<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 well written and error free. A minor issue is that, at times, the font is too small to easily read the text (as an example, in one chart, Table 2 Evaluation Questions on p 23).	
<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>	No	The report is 100 pages, significantly exceeding the 70 page limit for CPEs.	
<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 manner with clear distinctions between sections.	
<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>	Yes	The annexes include all required elements. It should be noted that Annexes 4 (sampling strategy), 5 (list of interventions) and 9 (list of persons met) are not listed in the table of contents although they are included in the annexes.	
<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	This serves as a stand-alone section and is an effective presentation of the main results.	
<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>	Yes	The summary is very clearly structured.	
<b>7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	Yes	The summary is concisely written and is 5 pages long.	
<b>2. Design and Methodology</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	Primary users are noted as being decision-makers in the UNFPA CO and within UNFPA as a whole, government counterparts, UNFPA executive board and other development partners.	
<b>2. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	There is a fairly lengthy section (12 pages) that describes country context, development challenges, progress towards achieving relevant internationally agreed development goals and external aid. A further 9 pages covers UNFPA's response and strategies.	
<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>	Partial	A simplified logic model for the CP is provided in table form in the main text. There is also a graphic depiction of the link between the CP outputs, the UNDAF outcomes, and national priorities. The adequacy of the existing logical model is not addressed and one issue that could have been raised is that some CP results currently considered outputs (i.e. strengthened national laws and policies) would generally be considered outcomes.	

<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 and how it was developed are described in the text. The full framework appears in the annex and includes the required elements.
5. Are the tools for data collection described and their choice justified?	Yes	The evaluators provide a clear description of how they used both qualitative and quantitative data sources, including through desk review, key informant interviews, focus group discussions, and observation of activities, as well as the rationale for each. They note that quantitative data was primarily obtained through secondary sources including the UNFPA Information System and Atlas System.
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)?	Yes	The mapping process is well described, including how stakeholders were grouped by UNFPA strategic plan outcomes/CPAP output and Atlas code. The map is presented in Annex 6 with supplementary information on previous and current Ministries of Armenia listed in Annex 7. It is explained that the preliminary findings, conclusions and recommendations were shared with UNFPA, and that the full draft report was shared with the Evaluation Reference Group and the Country Office.
7. Are the methods for analysis clearly described for all types of data?	Yes	The analysis methods are described. It is noted that data from interviews and FGDs was summarized on a standardized data summary sheet, and a semi-inductive thematic coding system was used. The quantitative data analysis was from secondary sources - national statistical data and the UNFPA country office M&E data.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The limitations are set out in a Risk and Risk Management table (table 3). Seven risks, as well as their likelihood, impact and mitigation measures are identified.
9. Is the sampling strategy described?	Partial	The sample was drawn from the stakeholder map to ensure all groups were represented in key informant interviews. Sites to be visited and participants for FGDs were based on consultations with UNFPA. It appears that the sampling was purposive (the report states 'deliberate' sampling was used to target the most vulnerable beneficiaries). While there was a risk and risk management table, detailing the limitations and mitigating measures, limitations to the sampling approach were not included.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The methodology enabled the collection of some forms of disaggregated data (i.e. the interview template has a line to enter stakeholder type but not gender of respondent; the focus group discussion guide did not include instructions for recording such information). Gender disaggregated data was not provided for the total number of evaluation respondents or for each type of data collection method. The List of Persons Met (Annex 9) included names with Mr. or Ms. prefix but did not include a tally of the numbers of males and females; nor was this information included in the main report.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	It is stated that gender and human rights principles were integrated throughout the evaluation process, and relevant evaluation questions and indicators were established. Beneficiary perspectives were obtained through focus group discussions, and there was an emphasis on examining programme benefits accruing to vulnerable and marginalized groups.
<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	There is a good explanation of how triangulation was employed, and it appears appropriate.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	There is a clear explanation of the sources of both types of data in the text. The specific documents and other sources are methodically cited in the evaluation matrix and throughout the findings. Both quantitative and qualitative data is used to support the findings. Reliability is referenced in respect to the lack of availability of key quantitative data.
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	The limited amount of primary quantitative data is noted as a limitation and the evaluators are explicit about how qualitative sources were used to address that gap.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	The evaluators reference the relevant UNEG and UNFPA guidance. They also mention that they strived for gender balance in selecting interview participants, had representation from vulnerable groups in focus groups, and attended to language and translation issues. The report could be strengthened by including disaggregated data on evaluation participants (by gender and stakeholder group), and by providing explicit information on how ethical considerations (such as confidentiality and informed consent) were taken up.

<b>4. Analysis and Findings</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To ensure sound analysis and credible findings</i>			
<b>1. Are the findings substantiated by evidence?</b>	Yes	Evidence is methodically provided for each finding. Visual aids, including text boxes that present best practices, are effectively used to highlight and supplement the evidence presented in the text.	
2. Is the basis for interpretations carefully described?	Yes	This is well done. The evaluators use the evaluation framework to guide their analysis; the framework included both indicators and assumptions for each question. These are then used in the findings, including with baseline/target/reported data for questions where that information is relevant.	
3. Is the analysis presented against the evaluation questions?	Yes	The analysis is presented against the evaluation questions.	
4. Is the analysis transparent about the sources and quality of data?	Yes	Sources of data are consistently referenced (whether documents, monitoring visits, key informant interviews). The report includes more than 200 footnoted citations of documents used. As mentioned above, the evaluators were upfront about the lack of primary quantitative data but were able to address this limitation.	
<b>5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</b>	Yes	Causal connections are provided throughout the findings. Examples include tables that show CP outputs and the link to the UNFPA strategic plan outcomes, as well as output indicators, targets, baseline and reported results information for each programme component. Unintended outcomes are taken up to some degree. Examples include how the Velvet Revolution has delayed CP implementation but has also opened new possibilities for UNFPA's work given the anticipated less conservative approach to governing (p 65), and the extent of success of UNFPA's work with non-traditional partners such as the military and clergy.	
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	This is done at several points. Under the SRH discussion, for example, the different outcomes for men and women are addressed. The discussion on Adolescents and Youth notes gaps in outreach to young boys at risk and to those with disabilities, and the need for more detailed capacity assessments of partner institutions.	
<b>7. Is the analysis presented against contextual factors?</b>	Yes	Contextual factors are used to explain the findings. For example, in assessing UNFPA support for regional emergency preparedness, the evaluators provide an overview paragraph on the risks Armenia faces from natural disasters and regional conflict, and also note that more information can be found in the country context section of the report. Another example is the way in the Velvet Revolution impacted how UNFPA worked with government (this is integrated into the discussion on effectiveness).	
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	Cross-cutting issues are addressed in the analysis. Examples include Finding 21 that looks at the success of the Gender Component in addressing the needs of the most marginalized groups (p 84), and an assessment of how support for cross-cutting issues may be more effectively delivered if the human resource archenteric were different (i.e. if a UNFPA staff member were housed with the UNFPA team rather than in the government ministry (p 68)).	
<b>5. Conclusions</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To assess the validity of conclusions</i>			
<b>1. Do the conclusions flow clearly from the findings?</b>	Yes	There are 10 conclusions and each specifies the respective evaluation question number(s) to which they respond/from which the findings were drawn.	
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Yes	The conclusion statements and succinctly stated supporting text are effective in conveying the essence of the findings, particularly given the length of the overall report.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no indication of bias.	

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	Each of the 9 recommendations include the number of the conclusion on which it was based.	
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	These are clearly presented and specify the target user. Although each has a subsection on operational implications, the text of this primarily focuses on an explanation of the recommendation and strategies for implementation but does not address the resource implications.	
3. Do recommendations appear balanced and impartial?	Yes	Both strategic and programmatic recommendations are provided, and they build on the strengths, opportunities, and challenges faced by the country programme. There is no indication of bias or partiality.	
4. Is a timeframe for implementation proposed?	Yes	It is noted that they are directed at the next country programme cycle.	
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	Each is given a priority rating (most are high). The ordering of recommendations could be revisited as the first is about placement of staff which does not appear to be as critical to highlight as other strategic recommendations. They are all presented in a way that enables a management response.	
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?		<p>a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3) There is not a specific objective but HRGE is clearly mainstreamed into the evaluation scope. = 3</p> <p>b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3) There was not a standalone criteria. HRGE was considered under Relevance and Effectiveness. = 3</p> <p>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) There is a question that looks at the extent to which the outputs contributed to mainstreaming of provisions to advance gender equality (EQ3) and one that looks at how policy advocacy and capacity building support has helped to ensure gender equality (EQ4). = 3</p> <p>d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?(Score: 0-3) This is addressed (there is a specific indicator for EQA 4 on the existence and availability of gender disaggregated data, and this is used in the analysis) and shortcomings were identified. Indicators = 3</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3) Although the evaluators note that they aimed for gender balance, they were not explicit about how the evaluation was gender responsive and evaluation participants were not gender disaggregated. = 1</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) Mixed methods and a participatory approach were used. = 3</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Data sources were diverse and well cited, and triangulation was evident. = 3</p> <p>d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3) Evaluation participants included a good range of stakeholders, including vulnerable groups. = 3</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) Guidance on ethical standards was cited as informing the methodology, however there were no details on how this was done = 1</p>	

