

Organizational unit: Egypt Country Office **Year of report:** 2017

Title of evaluation report: Evaluation of the UNFPA 9th Country Programme in Egypt (Mid 2013-2017)

Overall quality of report: Good **Date of assessment:** 21 September 2017

Overall comments: The report is well written and well structured. The approach to data collection is sound and the analysis that followed is of good quality. The evaluators noted the limitations to the methodology and proposed mitigating steps to reduce their impact on the quality of the findings. Lessons learned are described in a separate section. Findings are grounded in well-triangulated data/well-substantiated: there are references to the documents, interviews, and other sources of information - and the basis for interpretation in the analysis is carefully described. All the conclusions are linked to the findings and all the recommendations follow from the conclusions. The report has too many recommendations though and does not identify their human and financial cost implications. Evaluation objectives included gender and issues related to gender based violence.

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
--	---	--------------------------------------	--	---------------------------	--	--	--	---

Quality Assessment Criteria	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>		
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Fair
<i>To ensure the report is comprehensive and user-friendly</i>			
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?	Yes	The report is easy to read and understand . It is written in an accessible language with minimal grammatical errors.	
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)	Yes	The report is of a reasonable length (68 pages).	
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)?	Yes	The report is structured according to the UNFPA guidelines. A clear distinction between the analysis/findings , conclusions recommendations and lessons learned sections, has been made	
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?	Yes	The annexes are comprehensive. They include the TORs, methodology, bibliography, evaluation matrix, list of interviewees, interview guides, etc.	
<i>Executive summary</i>			
5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?	Yes	The executive summary is included as a stand-alone section and presents the main evaluation results.	
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)?	Yes	There is a clear structure for the executive summary. All the main sections of the reports , including the purpose, the objectives, methodology, findings, conclusions , recommendations and lessons learned are included.	
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The executive summary is concise but a a bit too long (slightly over 7 pages).	

2. Design and Methodology	Yes No Partial	Assessment Level:	Fair
<i>To ensure that the evaluation is put within its context</i>			
1. Does the evaluation describe the target audience for the evaluation?	Yes	The report does identify the target audience for this evaluation. They include: UNFPA senior management, executive board, donors, national partners, and CP managers, the Government of Egypt (GOE).	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The country context is well explained in chapter 2, including development challenges and country strategies.	
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes	The intervention logic has been mapped out and described in the report.	
<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 designed in accordance with UNFPA requirements. The evaluation questions are clear and correspond to the OECD–DAC evaluation criteria, including CO strategic positioning with regard to responsiveness, added value and coordination within UN country team in Egypt. The evaluation framework is clearly described in the text and in the evaluation matrix. The evaluation matrix is found in annex 18, and is designed in accordance with the UNFPA Handbook on Conducting a CPE. The evaluation matrix includes the evaluation questions, indicators, data sources, methods for data collection. It does not include any assumptions though but the information can be found in the text.	
5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection have been described and their choice justified.	
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	All the stakeholders have been identified, with each's role during data collection explained. The report explains that “Team consultations on findings received from interviewees were consulted further for data analysis, synthesis and reporting” (p. 5).	
7. Are the methods for analysis clearly described for all types of data?	Partial	The methods for analysis are not clearly described in the methodology (though data collection methods are).	
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The report acknowledges methodological limitations and how they affect the evaluation. The consultants had limited time available for primary data collection, challenges in assessing activities that were implemented between 2013 and 2016, and they mentioned “halo bias” (pp. 114-115). There was no evidence, however, of bias.	
9. Is the sampling strategy described?	Yes	The sampling strategy is described: sampling was illustrative, based on the stakeholder mapping provided by the UNFPA CO and “snow-balling” exercise (p. 119, p. 122).	
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology (document review, key Informant interviews, focus group discussions, site visits, online survey tool) enables the collection and analysis of disaggregated data.	
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Partial	The methodology is appropriate for assessing the cross-cutting issues as the consultants reached end line beneficiaries of services, including youth and women (Annexes 11 and 13). Cross-cutting issues were a part of the evaluation scope: “cross-cutting gender, poverty reduction, HIV prevention, YRH and the needs for vulnerable groups” (p. 17), though the consultants do not explain clearly their approach to assessing the cross-cutting issues.	

3. Reliability of Data	Yes No Partial	Assessment Level:	Good
<i>To ensure quality of data and robust data collection processes</i>			
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluators triangulated data collected in an appropriate fashion, as described in the methodology section and seen throughout the report. The evaluation consultants state that “Triangulation was ensured through systematic cross-checking of data and information sources on the one hand, and data collection tools, on the other hand” (p. xv). It is possible to find examples of triangulation in the text: “The tool proved to be a useful analytic tool and was funded by the CO to build a mixed method compliant with evaluation quality criteria and to triangulate FGDs and KPIs” (p. 121).	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation consultants clearly identified and used reliable qualitative and quantitative data sources. Most of the program partners, including the Government of Egypt, were used as sources of data (primary or secondary). The ERG, through comments, provided indications of the quality of data sources. Annex 9 provides detailed information on documents consulted.	
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 report does explain the limitations and risks associated with the data sources and gives an explanation of the actions that were taken to minimize their impacts. For example, the consultants mention “data gaps in acquiring detailed trainees/beneficiaries list” (p. 115).	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	The TOR specifies that “Team members will adhere and sign on the Ethical Guidelines for Evaluators in the UN system and the code of conduct” (Annex 1). The consultants underscored the fact that they tried to collect the data anonymously (p. 19, p. 113, p. 120). However, the methodology does not provide details on how the consultants addressed issues of discrimination and other ethical considerations during the evaluation.	

4. Analysis and Findings	Yes No Partial	Assessment Level:	Good
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	The findings are well substantiated by evidence: there are references to documents, interviews, and other sources of information (p. 11, p. 16, p. 29, etc).	
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretations is carefully described throughout the analysis by comparing evidence from different sources and discussing the quality of the data. For instance, the consultants explain that “The results showed that a network of peer youth education for RH was indeed established and youth capacities built, the effectiveness of this network needed further evaluation” (p. 13). Another example shows how the evaluators interpret the data: “However, the quality of target identification was a challenge. Y-Peer tried to expand their scope to target youth, adults, women and community leaders. However, such expansion did not match with the tools available by the team” (p. 31). In the meantime, there is more text than necessary and a lack of tables and diagrams in the section “Effectiveness.”	

3. Is the analysis presented against the evaluation questions?	Yes	The report presented the analysis against each evaluation question and went on to do additional work based on the online survey.
4. Is the analysis transparent about the sources and quality of data?	Yes	The analysis is transparent regarding the sources and quality of data. This includes the online survey which provided substance to a trend analysis.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluators assessed the changes that resulted from the implementation of the program activities and outputs, but the narrative text does not allow a review of cause-effect links easily. The following examples illustrate cause and effect links: <ul style="list-style-type: none"> • Interventions on “Capacity building of family planning” led to women “using different and proper contraceptives” (outcome) – p. 30. The consultants refer to the feedback from the beneficiaries and service providers to confirm the effectiveness of the intervention. • A documentary film for a model village supported by UNFPA (output) was used to raise awareness of local communities in improving population characteristics (outcome). Under the lessons learned section, the consultants noted unintended outcomes.
6. Does the analysis show different outcomes for different target groups, as relevant?	Partial	The analysis shows different outcomes for different target groups: UN agencies, country donors, central government, regional/decentral government, implementing partners, and end line beneficiaries of service like women and youth, but more analysis could have been done for youth, women and other socially vulnerable groups.
7. Is the analysis presented against contextual factors?	Yes	The analysis is presented against contextual factors which are well covered in the report: social, economic, and political challenges.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The consultants addressed cross-cutting issues throughout their analysis, for instance, vulnerable women in Cairo (p. 91), gender equality component (p. 32), human rights (pp. 23, 34, 48, etc).

5. Conclusions	Yes No Partial	Assessment Level:	Very good
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions flow directly 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?	Yes	The consultants provide thorough explanation of the underlying issues of the programme like “no analytical lessons learnt sufficiently drawn within Cycle... Spot check audits are not a replacement to capacity-related aspects such as bureaucratic delays or blockade... lack of direct intervention and capacity building activities based on robust data collected” (p. 57).	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	Conclusions appear to convey the evaluators' unbiased judgement: they are based on facts and examples from the findings. For instance, they evaluators explain that “During current CP, there is a tendency to collect useful empirical data to inform on programmatic targeting that is unutilized (e.g. DHS, SYPE)” (p. 58).	

6. Recommendations	Yes No Partial	Assessment Level:	Fair
---------------------------	----------------------	-------------------	-------------

To ensure the usefulness and clarity of recommendations		
1. Do recommendations flow logically from conclusions?	Yes	The recommendations are clearly connected to 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 clearly written, targeted at the intended users (addressee) and action-oriented. Operational implications are described, but not highlighted separately like in other UNFPA reports. For instance, the consultants explain that “The sequencing of SP, CPDs and CPAPs should be revisited, with SP arriving mid-cycle” (technical); “capacity investments and falls under frequent leadership changes” (financial) – p. 55. No information has been provided on their human and financial cost implications.
3. Do recommendations appear balanced and impartial?	Yes	The recommendations appear balanced and impartial across all program components.
4. Is a timeframe for implementation proposed?	Yes	The recommendations are categorized according to their urgency for implementation as long term, medium term and short term.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	Recommendations are clearly presented to facilitate appropriate management response and follow up on each specific recommendation. The recommendations are prioritized (Medium to High), but most recommendations have High priority. However, the number of recommendations (28) in the report is larger than usually desirable.

7. Gender	0 1 2 3	Assessment Level:	Good
------------------	------------------	-------------------	-------------

To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)		
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	The evaluation objectives included gender and GBV issues (p. 1). The evaluation matrix has GEEW indicators, for instance: “Policy dialogue is active on the related RH, PD, GBV topics” (p. 194), “National policies and mechanisms to combat GBV, in particular FGM/C, in place, and adopted by the relevant institutions” (p. 200) and others.
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 criteria are based on OECD-DAC criteria: relevance, effectiveness, efficiency, and sustainability. Also, the evaluation focused on CO strategic positioning with regard to responsiveness, added value and coordination with UN country team (p. 109). Therefore, the evaluation criteria do not address GEEW specifically. The evaluation questions have gender component: 1) women and young people; 3c) gender-based violence against women and girls; 3d) gender equality; 3f) GBV in Egypt (Annex 18).

<p>3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?</p>	<p>2</p>	<p>The methodology explains that “The anonymous tool enabled the surveyed respond freely and useful insights were gained...” (p. 123). The interview guides include instructions for the consultants to “Stress the confidentiality of the sources or the information collected” (p. 125).</p> <p>The methodology discusses that “The evaluation will adopt an inclusive approach that involves a range of stakeholders to generate diverse views on the program performance” (p. 71). Even though the consultants do not describe gender-responsive evaluation methods/tools and data analysis techniques, we can find examples of women involvement in the evaluation in the text “This figure is inclusive of 3 female meetings under the humanitarian component” (p. 5). Desk review and data analysis stages included gender analysis which is evident from the annex 9 “Documents Consulted” and the Annex 13 “Classification of Interviewees.” For instance, 29 Syrian women were interviewed (p. 122).</p>
<p>4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	<p>3</p>	<p>The evaluation findings include gender-responsive data in assessing all program components. Reproductive Health implies interventions for empowering women (Output 1), women of reproductive age (Output 2); Gender Equality - women of reproductive age (Output 2), gender-based violence against women and girls (Output 3); Population Dynamics - women of reproductive age (Output 2) – pp. 80-81. The data on the achievement of the outputs is presented in the Findings (Chapter 4) and the Evaluation Matrix (Annex 18). The evaluation conclusions reflect a gender analysis: C3 – “CO is an organization with relevant mandate for Egypt and focused on RH, GE...” (p. 56); C12 - “a success story for community-based interventions that empowers women...” (p. 59); C 17 - “While strong contributions to protection SGBV mechanisms are notable through the JP FGM, there are no significant contributions to active GBV response mechanisms in place” (p. 60), including C 18, C 19, and C 27.</p> <p>The evaluation recommendations reflect a gender analysis: R6 – “10th Cycle programming should revolve around a single outcome relating to RH with sub-components to GE, PD, FP, MH, and humanitarian response” (p. 57); R9 – “...in-depth analysis of statistics compiled and collated by CO to inform RH, GE and PD policy making” (p. 58); including R17 and R19.</p>

(*) 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		
Total scoring points	11	58	31	

Overall assessment level of evaluation report		Good		
	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?

- What aspects to be cautious about?

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

The data have been obtained reliably and the conclusions flow from this. While some recommendations could have been more precise, on the whole they provide a good basis for preparing the next country programme.

Consideration of significant constraints

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

- Yes No

If yes, please explain: