

Organizational unit:	UNFPA	Year of report:	2019
Title of evaluation report:	Government of South Africa/UNFPA 4th Country Programme 2013-2019 Evaluation		
Overall quality of report:	Good	Date of assessment:	11 December 2019

Overall comments: This report is a comprehensive evaluation of the UNFPA South Africa 4th Country Program, and the overall rating is good. The strengths of the evaluation are in the reliability of data, analysis and findings, integration of gender equality and human rights, and conclusions. The overall evaluation report is written in a logical manner and is well structured. The findings are based on the data and grounded in a clear contextual analyse. They are well-balanced between those that are positive and those that indicate areas for improvement. The evaluation findings, conclusions and recommendation reflect a gender analysis. However, the evaluation's methodology and recommendations could be improved. Although concise and action-oriented, the recommendations lack targets on implementation and their time-frame and resource implications are not clear. The methodology section could also be improved by providing a more detailed description of how sampling was done, how it conformed to ethical guidelines, how data was collected and analyzed, and disaggregated.

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:	Very good
<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	Yes, the report is easy to read and understand. There are a few grammatical errors, but they are not significant to understanding.	
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 main report is 61 pages without the executive summary, 66 with the summary.	
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 in a logical way, following the recommended structure, with a clear distinction between sections. The report is generally user-friendly, comprehensive, and written in accordance with OECD-DAC evaluation criteria.	
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 had all required information. While a specific annex describing the stakeholder consultation process was not included, the process was described in the text (p.14).	

<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 provides an overview of the evaluation, and is well written as a stand-alone section including key elements of the evaluation, such as objectives, methodology and conclusions and recommendations.	
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	All of the key elements are found in the executive summary.	
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	At five pages, the executive summary is within required page limits. The findings are structured by the evaluation criteria similar to the overall report format. The lessons learned, conclusions, and recommendations were concise.	
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 target audience was described in the executive summary and the main report as "The primary users of the CPE are the decision-makers within UNFPA and the Executive Board, government counterparts in South Africa, and other development partners including donors, the civil society, the private sector, as well as other UN agencies." (p. 13).	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	Chapter 2 clearly outlines the country context in terms of socio-political and economic contexts, challenges and national responses for sexual and reproductive health (SRH), adolescents and young people, population dynamics, gender equality and women's empowerment. It also includes a brief summary of the role of external assistance. Chapter 3 describes UNFPA's response and programme strategies. The chapter covered the institutional context, specifically the evolution of the country programme, its geographic coverage as well as partnering agencies.	
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	The evaluation report provides a theory of change in Figure 1 on page 12, but there is no description of the theory of change to help the reader understand the linkages between inputs, outputs, and results, and the assumptions and risks underpinning it. Section 1.3 on the "Reconstruction of the Theory of Change" could be strengthened with a narrative summary of the key components of the theory of change, namely a more thorough description of the components and their linkages.	
<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 can be inferred from the methodology, and the report provided a good description of the overall evaluation process starting with document review to stakeholder validation of the evaluation findings and dissemination. The evaluation matrix has all necessary components including the evaluation questions, assumptions, indicators, data sources and methods for data collection.	
5. Are the tools for data collection described and their choice justified?	Partial	The data collections tools are described. The template for the key informant interviews is provided in Annex 5. The report methodology section could be strengthened with additional description to show how data was collected (in addition to a list of the data collection tools used). The evaluation does not appear to provide clear justification of the methods (why they were selected) although the evaluation report does state that the purpose for document review was to understand the design of the programme and its underpinning theory of change.	
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)?	Partial	While there is no formal stakeholder map, the evaluators clearly specify the groups of stakeholders consulted. The overall stakeholder consultation process is not fully described, although the evaluation notes that the Evaluation Reference Group was central to the consultative process, a process which included consultations on the draft recommendations. The evaluators provided a list of stakeholders consulted in Annex 2, although not disaggregated by sex.	
7. Are the methods for analysis clearly described for all types of data?	Yes	The evaluation report mentions deductively generating themes along the evaluation questions using qualitative data analysis. The methodology section could be strengthened by including a detailed description of the type of qualitative analyses used, although the evaluation team noted that they analyzed data against the indicators and assumptions for each evaluation question.	

8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The methodological limitations for qualitative data were acknowledged and mitigation efforts mentioned. The limitations related to change of staff, lack of some data in the earlier years of the CP implementation, gaps in secondary data, and unavailability of some key informants at the national level. Aside from mentioning the use of triangulation for addressing methodological limitations, the process was not fully described beside identification of convergence and some divergence of the responses for each evaluation questions. The report could be strengthened by descriptions of how triangulation was done, possibly by the use of some form of visual illustrations.
9. Is the sampling strategy described?	Yes	The evaluation team reported using a purposive sampling approach to select key informants for the e-questionnaire to collect additional data on gender mainstreaming and human rights approach to programming. The criteria for the selection of site visits was based on where UNFPA was working. How interviewees were selected in terms of factors determining who to interview is also described. However, any limitations to this approach (and mitigating actions) were not mentioned.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The evaluation consultants used a mixed methods design, but they do not describe how this approach enabled the collection and analysis of disaggregated data nor reference methods and techniques they used for analysing disaggregated data (though a table detailing the range of stakeholder groups consulted is provided).
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	While the evaluation report does not explicitly include details on the appropriateness of the evaluation methodology for assessing the cross-cutting issues, the evaluation team used an e-questionnaire to collect additional data on gender mainstreaming and human rights approach to programming. The team integrated gender in the analysis of the country context and throughout the evaluation findings, despite the initial lack of integration of a gender perspective in the evaluation design (a gender specialist joined the evaluation mid-way in the data collection phase).
3. Reliability of Data	Yes No Partial	Assessment Level: Very good
<i>To ensure quality of data and robust data collection processes are used in the findings</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Triangulation was achieved through use of multiple data sources and methods. However, the evaluation team could have provided a description of how triangulation was used to validate data and evaluation findings (for example a table or diagram could help the reader understand how triangulation was applied).
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluators were explicit about the source of qualitative data coming mainly from interviews and quantitative data from secondary sources and the e-questionnaire. The evaluation team provided footnotes to track evidence from its source to use and interpretation. Nevertheless, it would have been useful for the team to describe how they ensured the quality of the qualitative data. The team noted a lack of quantitative data to back certain qualitative findings such as the reduction in maternal deaths through the establishment of Obstetric and Neonatal Ambulance (ONA) service (p. 42).
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 team reported mitigating occasions of potential data gaps through interviewing and seeking clarification from UNFPA staff and stakeholder and conducting validation workshops with stakeholders. However, the evaluation would have benefited from a more explicit discussion of bias introduced from data sources and how this was mitigated (for example, any potential bias from UNFPA staff involvement in the selection of persons to be interviewed - and mitigation strategies used).
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	There is no evidence in the presentation that there were issues of discrimination.
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 evaluators substantiated findings with evidence throughout, evidenced by several footnotes. The evaluation could have been further strengthened by substantiating evaluation findings further using direct quotes from the qualitative data.
2. Is the basis for interpretations carefully described?	Yes	The evaluation team noted using indicators and assumptions for each of the evaluation questions in the evaluation matrix for interpretation. For example, they reported that "Data from each source (documents and each category of key informants) was analysed against the indicators and assumptions of each evaluation question to identify the responses to each evaluation question from each data source" (p. 15).
3. Is the analysis presented against the evaluation questions?	Yes	In each case, the analysis is presented by evaluation question. The relevant evaluation questions are listed at the beginning of the discussion of each criteria, and the analysis is structured accordingly. In each question relating to the main areas (relevance, effectiveness, efficiency, sustainability and coordination), there is a summary of findings and then, for each finding, a detailed description of the basis for the finding.
4. Is the analysis transparent about the sources and quality of data?	Yes	The evaluation team were transparent about sources of data, through, for example, the use of footnotes throughout the report. They reported consulting with UNFPA staff and stakeholders whenever they experienced data gaps. The evaluation team noted throughout the discussions on findings the instances of lack of quantitative data. For example, the team noted that the reporting of the integration of gender and human rights-based approach lacked gender disaggregated data ("Reporting on the activities of this programmatic areas does include gender disaggregated data" (p. 51).
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The framework for assessing causal connections was well established in the evaluation matrix, and the evaluators were careful to show the theoretical links between the intervention (activities and outputs) and explanation of end results (intended outcomes). The effectiveness section is very sound. The evaluation delineated in tabular form and elaborated in the narrative text summary achievements of each outcome of the country programme and relevant outputs. While a direct cause-effect linkage between outputs and outcomes was not established as such, the report provided specific improvements that can be made to interventions. In a few cases, unanticipated outcomes, particularly in terms of far greater than expected results were reported. The report also included some negative unintended consequences of the interventions, such as trained personnel moving on to better paying jobs upon completion of training. The report occasionally provided non-achievement of outcomes due to either changes in the program contexts or gaps in programme design. For example, the evaluation team noted on page 35 that "the CP did not take deliberate steps to align to the Sustainable Development Goals (SDGs) which came into action after 2015."
6. Does the analysis show different outcomes for different target groups, as relevant?	Partial	Yes, the analysis examines outcomes for the different target groups in the relevant sections of the report such as adolescents and youth. However, the report does not clearly disaggregate results/outcomes by different stakeholder groups.
7. Is the analysis presented against contextual factors?	Yes	The report provided a detailed and lengthy section on contextual factors - in fact nine pages of text and statistical information (pp. 18 - 26) prior to presenting UNFPA response and programme strategies in the 4th CP. For example, the evaluation team reported that the development of the 5th CP was postponed to 2020 at the request of Government so that it would be harmonized with the Government and UN Strategic Cooperation Framework planning cycles. Another example is how the capacity for coordination of GBV was weakened by leadership changes in the Department of Women, Children and Persons with Disabilities (DOWCD).
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The evaluation report integrated cross-cutting issues such as gender equality and human rights throughout the report. Furthermore, the report provided summary analysis how the CP integrated gender and human rights-based approach in the different program outcomes. For example, the evaluation reported that the 4th CP is aligned to the key policy and legal framework for gender equality and women empowerment in the country: (i) the South Africa's National Policy Framework for Women's Empowerment and Gender Equality (NPFWEG), (ii) the Domestic Violence Act, (iii) the National Plan of Action for Children (NPAC) South Africa 2012-2017, and Integrated Programme of Action on Violence Against Women and Children 2013-2018. Furthermore, the evaluation reported that "All the SRH and HIV interventions supported by the CP addressed the SRH needs of adolescents and young women. However, there was a gap in the involvement of men in maternal and child health and family planning services" (p. 45). Finally, the evaluation reported noted that the 4th CP addressed underlying vulnerabilities of girls and women to gender-based violence which showed a demonstration of a rights-based analysis and approach to programming. The evaluation reported that the 4th CP applied gender equality and human rights approach across all outcome areas although this was not adequately integrated into targets, program design, implementation, and reporting.

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 from the findings and follow a clear structure, divided into strategic and programmatic focus. The evaluators specified the respective evaluation question numbers that are linked to each conclusion.	
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 evaluation team provided conclusions and recommendations at both the strategic and programmatic levels, thereby indicating their understanding of the underlying issues of the CP. The theory-based approach used in this evaluation provided a framework for assessing the extent to which the observed differences/results were a consequence of the 4th CP implementation in tandem with the complex operational environment.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no indication of bias. The conclusions reflect key findings and the validation workshops with key stakeholders was appropriate for checking potential biased judgment by the evaluators.	
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	Yes, it is clear that the recommendations logically flow from the findings and conclusions. Each recommendation refers to the specific conclusion from which it is derived.	
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 clear and action-oriented, although the report does not explicitly state and define the intended users or implementer of each recommendation, except recommendation 7a that is explicit to UNFPA. Furthermore, there is no specific reference made to the human and financial implications.	
3. Do recommendations appear balanced and impartial?	Yes	The recommendations do not have any evidence of bias.	
4. Is a timeframe for implementation proposed?	Yes	The timeframe for almost all is the time during which the 5th Country Programme is being developed.	
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Partial	Recommendations are not prioritized as high, medium and low per UNEG standards for evaluations and per the terms of reference. However, they are clear and can be used for the 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?		While the evaluation did not include a specific objective dedicated to gender equality, the evaluation appears to have mainstreamed gender equality in the application of the criteria and the evaluation developed five questions to ensure that the analysis assessed how gender equality and human rights was mainstreamed across the country programme, noting that “the analysis of gender and human rights was guided by the following questions (drawn from the UNEG gender framework). 1) Relevance: To what extent are interventions aligned with international gender instruments, national policies on gender and the different needs of men and women? To what extent are the interventions informed by substantive gender analyses that identify underlying causes and barriers gender equality? 2) Effectiveness: To what extent does the Theory of Change and results framework of the intervention integrate gender equality? To what extent was a gender mainstreaming strategy incorporated in the design and implementation of the interventions? 3) Efficiency: Does the short-term process achievements (participation and inclusiveness, etc.) and medium-term results (developing an enabling environment, building capacity, etc.) integrate and reflect gender equality? Was provision made for adequate resources for integrating gender equality in the CP interventions? 4) Sustainability: Has institutional change conducive to systematically addressing gender equality concerns been created? And 5) Coordination: To what extent has UNFPA proactively driven and supported the meaningful integration of gender equality across interventions?” As such, the evaluation was able to capture a fragmented approach to gender mainstreaming across the programme. Finally, the evaluation noted that insufficient information was collected during the programme implementation period to measure progress on human rights and gender equality, especially data on alignment of the 4th Country Programme to UNFPA Strategy and the achievement of planned results.		
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		While evaluation questions reflected gender and human rights, the evaluation methodology does not specify how gender issues are addressed throughout the evaluation, and could have provided a description of how gender and human rights were considered in the data collection, sample size/make-up, description of the locations of the interviews, and data disaggregated by sex. Citations in the footnotes indicate that data was triangulated, though it is not fully clear how triangulation and validation guaranteed the inclusion of a diverse range of voices including those of rights holders/beneficiaries. Moreover, it is not clear how the evaluation methods and sampling frame addressed diversity of stakeholders affected by the intervention without a clear comprehensive stakeholder mapping/analysis. The interview and focus group protocols did not indicate how data complied with ethical guidelines for informed consent and respect for confidentiality.		
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		While the findings included data analysis that explicitly triangulated quantitative and qualitative data based on data sources cited in footnotes, there was minimal use of quotes or stories to highlight the voices of different groups. The quantitative data is also not disaggregated by gender, although the evaluation report notes that the programme did not collect sufficient gender disaggregated data. While the evaluation does discuss some negative unintended consequences of interventions generally, the evaluation report does not provide an explicit descriptions of unintended effects of the country programme on human rights and gender equality. The evaluation provides three specific conclusions and four recommendations explicitly addressing gender equality, particularly focused on gender-based violence.		
<p>(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. 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).</p> <p>(**) Scoring uses a four point scale (0-3).</p> <p>0 = Not at all integrated. Applies when none of the elements under a criterion are met.</p> <p>1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.</p> <p>2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.</p> <p>3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.</p> <p style="text-align: center;">Overall Evaluation Quality Assessment</p>				
	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		
Total scoring points	29	47	24	
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 evaluation uses tools and concepts well to produce a good evaluation.

Consideration of significant constraints

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

Yes No

If yes, please explain: