

Organizational unit: Year of report: **2022**

Title of evaluation report: **United Nations Population Fund (UNFPA) Nigeria 8th Country Programme (2018 – 2022)**

Overall quality of report: **Very good** Date of assessment: **5 May 2022**

Overall comments: The evaluation provides a strong context and analysis of the ambitious CPE. The findings are presented by criteria and follow the logic of evaluation questions, providing an appropriate structure for the intended audience. The evaluation is notable for its thorough analysis of the ToC. The methodology enabled the participation of a range of stakeholders and is mostly well described although it would be more comprehensive if it included a stakeholder map, a discussion on the selection of the sample of rights holders/beneficiaries, and identification of the total number of people reached by each data collection method. Findings are thorough, evidence-based and triangulated although in the main report evaluators could have been more specific about which document sources were used for the effectiveness criteria and could have brought out the perspectives of the range of stakeholders more clearly, especially those of vulnerable groups. The evaluation provides concise conclusions that are linked to each evaluation question. Recommendations are also specific and useful to improve Nigeria's next programme cycle. The 200 pages of annexes, provided in a separate document, are more bulky than usual and include a somewhat overly detailed evaluation matrix and list of atlas projects.

Disability inclusion is reflected in the evaluation scope, questions, under multiple criteria within findings and in one recommendation. It is particularly well assessed under Effectiveness, including in a discussion on access rates by the CO to build capacity such as the appointment of a focal point on disability. However, it is not clear if people with disabilities were assessed or mentioned in the

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 *Insert assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)*

1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Very good
--	----------------------	-------------------	------------------

To ensure the report is comprehensive and user-friendly

1. Is the report structured in a logical way? 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? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Yes	Overall, the report is well structured, following an expected logic. It is easy to read with minimal grammatical/spelling errors (although some spacing issues occur, likely because the document was converted from a pdf file format). The language used is appropriate for the intended audience (UNFPA CO and sub-offices, among other partners and stakeholders). The report consists of distinct chapters for findings, conclusions and recommendations. In the future, evaluators should be encouraged to use rights based language.
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 reasonable in length for a CPE at 70 pages, including the executive summary.
3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?	Yes	The annexes include all of the required elements and are presented in a separate file. In addition to those required, the annexes include a very lengthy list of Atlas projects of CP by outcome area (which is over 50 pages long), a CP Achievement Table, and a table on budget implementation rates.

Executive summary

4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) intended audience; iv) Methodology; v) Main results; vi) Conclusions and Recommendations?	Partial	The executive summary is a well written stand-alone section; it includes purpose and objectives as well as intended audience and methodology of the CPE. The latter does not specify however the number of evaluation participants (from semi-structured interviews, FGI and site visits) which would be useful information for understanding the breadth of the evaluation process. Although it does not presents main findings separately, main conclusions provide detailed results. Therefore, conclusions are sufficient. Recommendations are also provided.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is carefully and succinctly written. It is four pages long.

2. Design and Methodology	Yes No Partial	Assessment Level:	Very good
----------------------------------	----------------------	-------------------	------------------

To ensure that the evaluation is put within its context

1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The report includes development challenges as well as national strategies, providing a broad context in order to understand the implementation of the CPE. It also details the demographic variables (in particular the population growth) which are helpful to understand the development context.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The report includes a subsection (1.3.2) within the methodology section, in which the overarching ToC is analysed in detail and suggestions for improvements are made. The results chain logic for each thematic area is also assessed in Findings. Graphic depictions of the ToCs are provided in Annex B.

To ensure a rigorous design and methodology

<p>3. 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?</p>	Yes	The evaluation matrix included in the annexes is complete with assumptions to be assessed, indicators, sources of information and methods and tools for data collection. Each evaluation question in the matrix includes highly detailed main findings as well, providing a robust evaluation matrix in order to understand the CPE.
<p>4. Are the tools for data collection described and their choice justified?</p>	Yes	Data collection tools are described and their justification/rationale is provided in the main report and the evaluation matrix. They include: Key Informant Interviews and Focus Group Interviews (FGIs) for primary data collection and document review for secondary data collection, as well as an extensive number of site visits (27).
<p>5. 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)?</p>	Partial	The report does not include a stakeholder map (nor is it included in the annexes) even though the ToR explicitly asks for one to be presented. Some insight is provided in Annex I Implementation Rates which lists implementing partners by year as well as their budget and budget utilization rate. Rights holders participated in FGIs and from the annexed list of persons consulted it is evident that these included GBV survivors and youth. However, the text in the report could have been more clear about the extent to which FGIs included vulnerable and marginalized groups, including those with disabilities. Table I.3 indicates that the draft was reviewed by the CO and ERG, and that their comments were incorporated into the final document. As such, they would have had a chance to comment on the recommendations.
<p>6. Are the methods for analysis clearly described for all types of data?</p>	Partial	It is mentioned that contribution and trend analysis process were employed. However, these are described in general terms. It would have been useful to explain the mechanical aspects such as coding processes and any software applications used.
<p>7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?</p>	Yes	Limitations encountered, particularly in stakeholder sampling and field work logistics, are described as well as the mitigation strategies taken to minimize the obstacles that the evaluation team faced.
<p>8. Is the sampling strategy described?</p>	Partial	In line with the participatory methodology, the sampling strategy description includes purposive stakeholder selection based on discussion between the evaluation team and the evaluation manager and technical thematic leads, mentioning briefly the criteria considered in their selection. However it is not clear how beneficiaries were selected for FGIs or therefore how representative they were.
<p>9. Does the methodology enable the collection and analysis of disaggregated data?</p>	Partial	Although the methodology allows the disaggregation by gender of the stakeholders consulted, neither the table summarizing stakeholder consultation (Table I.2) in the report nor the list of annexes presents data disaggregated by gender or by ability/disability.
<p>10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?</p>	Yes	The evaluation design methodology appears appropriate to assess the extent of integration of cross-cutting issues into the CP, including the use of rights-based approaches and the consideration of the specific vulnerabilities of GBV survivors and persons with disabilities. Questions in the evaluation matrix include cross-cutting issues (for example: 1.b. refers to priorities including vulnerable or marginalized groups; 2.b. refers to the integration of human rights, gender perspectives and disability inclusion in the design, implementation and monitoring of the CP).
<p>3. Reliability of Data</p>		
<p>Yes No Partial</p>		<p>Assessment Level: Good</p>
<p><i>To ensure quality of data and robust data collection processes</i></p>		
<p>1. Did the evaluation triangulate data collected as appropriate?</p>	Yes	The evaluation findings triangulate data from both the primary and secondary sources/methods. The evaluation matrix shows triangulated data for each evaluation question.
<p>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</p>	Partial	The report appears to make adequate use of both quantitative and qualitative data, based on the sources of information defined in the evaluation matrix and accordingly included in the report. Nevertheless, reliability is not specifically discussed in the methodology section, although the findings appear to have appropriate use of qualitative and quantitative data.
<p>3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</p>	Yes	The methodology section notes the use of UNFPA Handbook 2019 and UN ethical standards, including UNFPA COVID-19 guidelines. It also mentions that confidentiality and privacy were assured throughout, and 'stakeholders being informed that they were free not to respond to any question they found uncomfortable'. These issues are also reflected in the data collection protocols.
<p>4. Analysis and Findings</p>		
<p>Yes No Partial</p>		<p>Assessment Level: Very good</p>
<p><i>To ensure sound analysis and credible findings</i></p>		

1. Are the findings substantiated by evidence?	Yes	Findings are comprehensive and based on diverse evidence from both primary and secondary sources. Although there is extensive citing of specific document sources in most sections, this is more vague under effectiveness where the source for secondary evidence mostly just mentions "document review". Throughout the findings, the primary sources are generally identified as "KI interviews" and/or "FGIs", and the specific stakeholder group is not provided in most cases. The evaluation matrix does provide more exact information so a reader looking for specific evidence would be required to read the findings and the matrix.
2. Is the basis for interpretations carefully described?	Yes	In all evaluation criteria, findings are supported by a thorough analysis of the evidence and well explained.
3. Is the analysis presented against the evaluation questions?	Yes	Findings are structured based on the evaluation questions, facilitating for the audience the main findings presented in a logical manner.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Overall the evaluation presents causal links between the interventions and its end results, based on the results framework of the CP. Analysis included contribution analysis, to assess how far the documented inputs and activities sufficed and were relevant to the outputs and likely to have contributed meaningfully to the intended outcomes. Under the effectiveness and coverage criteria, the report suggests unintended results, such as excessively punitive measures being legislated, including castration of convicted rapists, as an outcome of the domestication of the Violence Against Persons (Prohibition) Act. UNFPA advocacy activities to promote human rights and GEWE had involved this Act.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	Findings discuss the extent to which the CP reached different groups, i.e. including women, adolescents and youth, people and women with disabilities. For instance, the report mentions how UNFPA has supported states to strengthen capacities to prioritize adolescent and youth policies, health, and development. One finding reveals though that the level of training was inadequate for the depth of skill and understanding needed to support traumatized GBV survivors.
6. Is the analysis presented against contextual factors?	Yes	The analysis presented takes into consideration the context in which the CP developed, i.e. COVID-19 pandemic, among other factors.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The analysis includes cross-cutting issues of vulnerability, disability inclusion, gender equality and human rights. This was most evident under effectiveness and coverage in the discussion on sexual and reproductive rights.
5. Conclusions		
	Yes No Partial	Assessment Level: Good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	Conclusions are presented in a very succinct manner. Each conclusion is directly linked to the relevant evaluation questions and evaluation criteria.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Partial	Conclusions are systematized and presented based on the strategic plan outcomes themes, reflecting appropriate cross-cutting issues: strategic level; sexual and reproductive health services: adolescents and youth; gender equality and women's empowerment and population dynamics. However this section does not include anything on disability inclusion.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	Conclusions are presented in an unbiased manner, ensuring objectivity.
6. Recommendations		
	Yes No Partial	Assessment Level: Very good
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	Recommendations are directly linked to particular conclusions, therefore are derived in a logical manner.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	Each recommendation is targeted at different levels (although all of them are directed to the CO) and include a brief description of the operational implications the recommendations has, i.e. "need for adequate staffing".
3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?	Yes	The recommendations appear balanced and impartial and include cross-cutting issues.
4. Are the recommendations prioritized?	Yes	The priority level of each recommendation is explicit (high or medium).
7. Gender		
	0 1 2 3 (**)	Assessment Level: Very good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>		

<p>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>	<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) The assessment of cross-cutting issues is part of the scope of the evaluation, however although human rights, disability and internal displacement are some of the issues identified, gender is not. (1)</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) Gender equality was considered under Relevance & Responsiveness and Effectiveness & Coverage, with human rights mainly be considered under the latter. (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) EQ2 integrated specifically GEEW. (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) The evaluation does not assess the collection of disaggregated data to measure progress on gender equality and human rights, but it assesses the results indicators. (1)</p>
<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	<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) The methodology is not specific on how gender issues are addressed. Not all collected data is disaggregated by sex (only FGIs). (1)</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analysing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) The selected mixed-methods approach generally appeared appropriate, although sample size was not discussed. (2)</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) Contribution analysis was applied and appropriate triangulation guaranteed inclusion, accuracy and credibility. (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) The evaluation's participatory process included diverse stakeholders, however the universe could be more clear. Stakeholders included girls of reproductive age, sexual assault beneficiaries, GBV survivors, health workers, educators, among others. (2)</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) Ethical standards were referenced, however there is no specific ethical consideration mentioned regarding interactions with vulnerable groups/stakeholders. (1)</p>
<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3) The context section includes an adequate analysis on specific groups, such as adolescents girls and young women. (3)</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3) The findings provide disability and gender analysis and presents disaggregated data where possible, specially under the results framework. However, when citing stakeholders they are only referenced as "FGIs, KI interviews" so its unclear the weight given to voices of different groups. (2)</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) An unintended effect on implementation of advocacy to promote human rights and GEWE is included (legislation of excessively punitive measures). (3)</p> <p>d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3) The evaluation includes recommendations in regards to SRHR, GBV and human rights. (3)</p>

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted.

(**) Scoring uses a four point scale (0-3).

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	78	22		
Overall assessment level of evaluation report	Very good			
The evaluation integrates adequately cross cutting issues of gender equality, human rights and disability inclusion, even though is not included as part of the evaluation objective.				
Consideration of significant constraints (e.g. COVID-19 or civil unrest)				
The quality of this evaluation report has been hampered by exceptionally difficult circumstances: <input type="checkbox"/> Yes <input type="checkbox"/> No				
If yes, please explain:				