

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

**Title of evaluation report:** EVALUATION OF THE UNITED NATIONS POPULATION FUND (UNFPA) NAMIBIA 6th COUNTRY PROGRAMME (2019 – 2023)

**Overall quality of report:** Good **Date of assessment:** 13 March 2023

**Overall comments:** Overall, this is a solid evaluation report that can be used by decision-makers with confidence. The evaluators were very thorough in documenting the range of country programme activities and data obtained from the evaluation. They were also meticulous in the level of detail provided in the evaluation matrix. The evaluation design was based on the Theory of Change and the results framework. The mixed methods approach enabled adequate triangulation of evaluative data in addition to validation through stakeholder consultations. The evaluation findings systematically responded to the evaluation questions in a balanced way, with rigour and extensive use of qualitative data. The factors facilitating and/or hindering the achievement of results were reported. Furthermore, the findings, conclusions, and recommendations fully integrated human rights, gender equality, women's empowerment, and disability. The recommendations are targeted and actionable with very clear prioritization. Despite these strengths, some shortcomings were noted. The methodology section could have been more clear about the approach to sampling and, in particular, the extent to which the programme participants involved in FGDs were representative of vulnerable populations. The findings were very descriptive; the inclusion of key findings statements and more sub-headings/bolded text would have been useful for making the basis of the analysis more apparent. In addition, the sources of evidence could have been more specific by citing specific documents and stakeholder groups.

The evaluation was disability inclusive in that persons with disabilities and representative organizations were included as respondents (this was evident in the annexed evaluation matrix and the list of participants), and the evaluation questions and analysis specifically addressed the extent to which the CP was disability inclusive. There were conclusions on disability although the recommendations did not explicitly address this. It would have been useful for the methodology section to include an explanation of how persons with disabilities were involved and any accommodations made to support their participation.

**Assessment Levels**

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

<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level;	<b>Fair</b>
--	----------------------	-------------------	-------------

*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)?	Partial	The report is logically structured overall, although the discussions on Challenges and on Lessons Learned appear under 4.7 Connectedness, instead of being separate sections. Otherwise, all the sections of the report are well-labelled. There is some use of tables and graphs, although more visual aids and formatting techniques would be useful to reduce the density of text in several section. The report would also have benefitted from close final proofreading to address minor grammatical errors and typos, and the variation in typeface.
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)	Partial	At 75 pages excluding the annexes, the report is longer than the 70-page requirements for CPEs.
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 report contains the necessary annexes, including the TOR, the evaluation matrix, reconstructed theory of change, data collection tools, stakeholders map, list of consulted stakeholders, and references for documents consulted.

*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 reads as a stand-alone document and includes all of the required sub-sections. However, the recommendations section is presented in paragraph form and is lengthy; the use of numbering and/or bolded topic areas would make the recommendations more clear.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The executive summary is 6.5 pages in length. It could have been more concise by combining finding and conclusions as the latter extends over 2 pages and does not consistently provide a higher-level overview.

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level;	<b>Good</b>
----------------------------------	----------------------	-------------------	-------------

*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	Chapter Two clearly describes the country context of Namibia including the developmental and socio-economic situation regarding sexual and reproductive health rights and adolescents and youth, gender equality and women empowerment (GEWE), population dynamics, and the national development strategies and the role of external assistance.
--	-----	--

<p><b>2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?</b></p>	<p>Partial</p>	<p>The report clearly outlines the two major outcomes and two thematic areas of programming with distinct outputs of the sixth UNFPA Country Programme in Namibia 2019-2023 based on the results framework in Table 3 and TOC. In addition, the TOC in Annex 2 presents the evaluator's re-construction of the interventions' logic, depicting the results chain between outputs and outcomes, including assumptions and risks per outcome. Unfortunately, the images are blurry, making the content for each category difficult to read. It appears that an atypical approach was used in the re-construction with the outcome indicators feeding into the output indicators, and the outcome statements being framed as what the interventions will do (i.e., "strengthen capacity of institutions to . . .") instead of capturing the change envisioned as a result of the intervention.</p>
<p><i>To ensure a rigorous design and methodology</i></p>		
<p><b>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?</b></p>	<p>Yes</p>	<p>The evaluation criteria and evaluation questions are clearly outlined in Annex 3. The evaluation matrix is organised by the evaluation criteria, evaluation questions, the relevant indicators, assumptions, and methods for data collection and detailed lists of sources of information. In addition, the evaluation matrix includes detailed stakeholder information and key findings for each evaluation criteria. The matrix is also referenced in the main part of the report.</p>
<p><b>4. Are the tools for data collection described and their choice justified?</b></p>	<p>Partial</p>	<p>Evaluators provide detailed justification for the chosen evaluation design. Tools included a literature review, key informant individual interviews, and FGD. The data collection tools are presented in Annex 4. However, the evaluators did not include an observation protocol although direct observation is mentioned as one of the main data collection processes.</p>
<p><b>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)?</b></p>	<p>Partial</p>	<p>Annex 5 provides a table of UNFPA implementing partners, other partners, and right holders organized by outputs for each strategic outcome. The report provides a description of stakeholder involvement at the different stages of the CPE including design, data collection, data analysis, and reporting especially at the recommendation formulation process, debriefing, and dissemination stages. It is stated that the evaluation reference group were consulted at each of the stages of the evaluation process, but the report does not state explicitly how vulnerable groups, such as persons with disabilities or living with HIV, or their representatives were involved in the evaluation or the extent they have been included in the sample.</p>
<p><b>6. Are the methods for analysis clearly described for all types of data?</b></p>	<p>Yes</p>	<p>The evaluators clearly described the data analyses methods as including qualitative content analysis and contribution analysis. It is stated that the contribution analysis assessed the extent to which the TOC contributed to the output and outcome indicators. Although the report does not describe the mechanical processes for how the analyses were conducted, the evaluators provided the definition of what these methods are and descriptions of how each method was applied.</p>
<p><b>7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?</b></p>	<p>Partial</p>	<p>Some limitations are stated such as the impact of short duration and scope of the evaluation limiting the collection of quantitative data and limited observation of actual effects of the CP interventions in the country, but the report could benefit from having a more structured discussion of the mitigation of the methodological limitations. Data gaps, which are noted in various places in the findings, could also have been highlighted as a limitation.</p>
<p><b>8. Is the sampling strategy described?</b></p>	<p>Partial</p>	<p>The sampling process was described in general terms in section 1.3.4 of the report. However, although the universe from which the sample was drawn was described as the list of major stakeholder categories of the CP6 outputs and outcomes, there is no mention of either the sampling criteria nor the sampling approach for how the stakeholders or sites visited were selected. It is not clear if the evaluators used stratified random sampling to ensure representation of the vulnerable population and what their desired sample was. It is stated that the evaluators collected primary data from an agreed sample of 132 stakeholders and the final sample is found in Annex 7 for the list of respondents, but the breakdown or overview of the sample is not provided in the methodology section. As such, the adequacy and representativeness of the sample is not readily apparent.</p>
<p><b>9. Does the methodology enable the collection and analysis of disaggregated data?</b></p>	<p>Partial</p>	<p>Although the TOR stated clearly that the evaluation should disaggregate data by sex, age, location, and other relevant dimensions, such as disability status the report only mentions disaggregating data along gender, but the other vulnerability criteria mentioned in the report were not provided. For example, Table 5 in Findings provides an assessment of the performance of the 6th GRN-UNFPA Country Programme, but there is lack of data by the different categories. With respect to primary data collection, it is not clear what proportion of men and women were interviewed in the agreed sample of 132 participants. The list of evaluation participants in Annex 7 does not clearly state the gender of all respondents.</p>
<p><b>10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?</b></p>	<p>Yes</p>	<p>Although the methodology section could be more clear on how the voices and perspectives of the range of stakeholders like vulnerable and marginalized groups were captured beyond having FGDs with beneficiaries, it is apparent from the evaluation matrix and list of evaluation respondents that vulnerable groups, including those with disabilities, were involved. In addition, the evaluation questions show the intent to collect and assess data on cross-cutting themes.</p>
<p><b>3. Reliability of Data</b></p>	<p>Yes No Partial</p>	<p>Assessment Level: <b>Good</b></p>
<p><i>To ensure quality of data and robust data collection processes</i></p>		
<p><b>1. Did the evaluation triangulate data collected as appropriate?</b></p>	<p>Partial</p>	<p>The evaluators emphasized that they triangulated data from different sources and methods to identify consistent topics, themes, and patterns, but there is no description of this process. It is not sufficient to simply mention that data was triangulated without providing a description (or showing more evidence) of how this was done to enhance validation of findings. Within the findings section, sources are mostly identified by data collection method (document review, KIs, FGDs) but are not specific about which documents and only infrequently identify the stakeholder group that is the source of qualitative data. Some gaps are also quite noticeable, for example on p. 48 where it is stated that teen pregnancy remains high, the only source of this statement is ET observations from visits to the regions.</p>

2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The report clearly states that sources for qualitative data were field observations, focus groups and interviews of evaluation participants and that quantitative data came primarily from secondary data. Data reliability is not explicitly discussed although evaluators do note using a solid mixed-methods approach with triangulation to ensure the credibility of the evaluation. They also highlight holding regular internal debriefing meetings to compare and validate data from interviews, and having validation meetings with the CO and ERG.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	Adherence to UNEG Ethical Guidelines is noted in the methodology section and ethical considerations - including informed consent, confidentiality anonymity - are evident in the data collection protocols. The evaluation matrix has questions and indicators that addressed issues of discrimination and there is evidence in the report that these issues were addressed throughout the evaluation.
4. Analysis and Findings	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Partial	The evaluation findings present a detailed narrative that responds to each evaluation question and appears to provide a balanced picture of the programme's strengths and weaknesses. The findings are supported by general references to document review, project monitoring data, interviews, as well as the inclusion of direct quotes from respondents - however, as mentioned above, in most cases the sources are only identified by the type of data collection method used. A further issue is that much of the content is descriptive, providing a lot of detail on CP activities, but could be more analytical. The inclusion of key findings statements would help to show how the data was answering the evaluation questions. At times, problematic issues are raised, such as "the ET found the rapid assessment of low quality" in discussing the readiness of post-GBV services (p. 37), but these are not further substantiated or discussed.
2. Is the basis for interpretations carefully described?	Partial	The findings included reference to the results framework of the 6th CP, the modified Theory of Change (TOC) presented in Annex 2, and the UNFPA global results framework as providing a framework for the interpretation of evaluation findings. The detailed account of the programme activities also aids in the interpretation. However, this section could have been more clearly organized, including by more use of sub-headings, to make it easier for the reader to see the basis for the findings. This is particularly the case for the Effectiveness discussion where having sub-headings for each output would have provided more structure to the analysis. It would also have been helpful if Table 5, which shows the outcomes and outputs and levels of achievement, was placed at the beginning of Effectiveness, rather than at the end, to give more context for the analysis.
3. Is the analysis presented against the evaluation questions?	Yes	The findings are presented by the evaluation questions and a brief summary of the findings for each question is given.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The findings are premised on the results framework and TOC in order to assess the causal linkages amongst output and outcome level changes. There is adequate attention to linkages particularly in the Effectiveness discussion where evaluators also note where there is still insufficient data on outputs to show outcome-level achievements. The Effectiveness evaluation question includes a component on unexpected results. Although the subsequent analysis does not explicitly address this, there is a thorough discussion on the covid response and some of the particular achievements and challenges during that time, as well as of innovative work done.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis has some reference to performance levels in the Effectiveness discussion however, beyond this and the discussions on disability, there is minimal attention to other target groups. The evaluators observed that the focus on the marginalized is broadly defined.
6. Is the analysis presented against contextual factors?	Yes	Context is consistently provided. Factors hindering and enabling the achievement of results are clearly discussed. The contextual factors that may have affected the delivery of programme outcomes are also presented in the reconstructed TOC.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Disability and women's empowerment are well mainstreamed into the analysis and there is also a specific subsection under Effectiveness on equity and vulnerability.
5. Conclusions	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions logically flow from the findings and context, and are presented at both strategic and programme levels. All conclusions are explicitly linked to the corresponding evaluation criteria in the findings and the corresponding recommendation.
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?	Yes	For the most part, conclusions provide a higher-level overview of findings. The underlying issues related to the country context, the UNFPA strategic directions, and the cross-cutting issues - including gender equality, equity and vulnerability, human rights, and disability inclusion - are addressed.
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:	Very good
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	Each of the 10 main recommendations are linked explicitly to its corresponding finding and the relevant evaluation conclusions. The recommendations are categorized as strategic or programmatic.	
<b>2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b>	Yes	All the recommendations are actionable and appear useful. Each recommendation clearly identifies the users responsible for implementation and propose clear operational implications/sub-recommendations for their implementation.	
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?	Partial	These appear to be impartial and reflect the strengths and weakness found through the evaluation process. There is some attention to cross-cutting themes, mainly including gender equality, equity and vulnerability although disability was not discussed.	
4. Are the recommendations prioritized?	Yes	The level of priority is identified for each recommendation ranging from intermediate to very high.	
7. Gender	0 1 2 3 (**)	Assessment Level:	Very 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?	3	<p><b>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)</b> Yes, the thematic scope includes GEWE as a cross-cutting theme. = 3</p> <p><b>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)</b> HRGE issues are mainstreamed into the criteria = 3</p> <p><b>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)</b> There is a dedicated question on the integration of GEEW under Effectiveness. = 3</p> <p><b>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)</b> The use of the results framework and indicators provided a framework assessing whether or not sufficient information was collected human rights and gender equality results. Data was disaggregated by gender, when possible. (3)</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	2	<p><b>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)</b> Evaluators note using the UNEG guidance on HRGE in evaluations but are not explicit about how the approach was gender-responsive. = 1</p> <p><b>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)</b> The mixed methods and participatory approach is appropriate for assessing GEEW, however the sampling process is not clear about the extent women's voices are incorporated. = 2</p> <p><b>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)</b> Although sampling is not well defined, evaluators state that the participatory approach ensured inclusion. Both triangulation and validation were evident. = 2</p> <p><b>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)</b> The sampling frame includes a wide range of stakeholders. However, it is unclear the extent to which particularly vulnerable groups were consulted since the report does not include the list of direct beneficiaries consulted during data collection and analysis phases.=1</p> <p><b>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3)</b> Ethical considerations are apparent but evaluators could have been more specific about their application to rightsholders = 2</p>	
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	<p><b>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)</b> The context section provides an appropriate analysis, including on GBV. = 3</p> <p><b>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)</b> The findings include a reasonably solid analysis of GE, however although the voices of those representing disability inclusion actors are apparent, the voices of women are not. = 2</p> <p><b>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)</b> The evaluation does not explicitly include the unanticipated effects of the intervention, although the discussion on covid response is somewhat relevant. =1</p> <p><b>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)</b> Both the strategic and programmatic recommendations specifically address GEWE issues. = 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			
<b>Total scoring points</b>	<b>29</b>	<b>64</b>	<b>7</b>	

**Overall assessment level of evaluation report** **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.

Gender equality and disability inclusion were adequately explored.

**Consideration of significant constraints (e.g. COVID-19 or civil unrest)**

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:  Yes  No

If yes, please explain: