

Organizational unit:	Tanzania CO	Year of report:	2022
Title of evaluation report:	Evaluation of the Government of the United Republic of Tanzania / United Nations Population Fund 8th Country Programme (2016/17 - 2021/22)		
Overall quality of report:	Good	Date of assessment:	20 April 2022
Overall comments:	<p>This report provides a detailed overview and analysis of the Country Programme. The intended methodology included document review, key informant interviews and focus group discussions with the sampling process being based on comprehensive stakeholder mapping. The annexed stakeholder table goes beyond most by indicating the relationship of each entity to the programme and a description of the partnership. The evaluation matrix is also well developed with detailed indicators and summarized key findings for each evaluation question. Although much of the methodology is well laid out, a major shortcoming is that the number of evaluation participants is not provided and as such, the representativeness of the final sample cannot be determined particularly regarding the extent to which rightsholders were able to be engaged. The evaluators could also have more fully explained the mitigation measure undertaken; the challenges are well described but it is unclear how they were resolved. The findings are typically substantiated by multiple lines of secondary/documentary and primary-level evidence; these are then linked to a large number of conclusions and recommendations than is typically seen. This is very detailed and quite text heavy report - although it clearly provides important insights for decision making, its usability would likely be increased if the presentation was more succinct and if there was more use of design/formatting techniques to make the document more accessible to the range of intended audiences.</p> <p>The evaluation performs strongly in respect to disability inclusion. Disability is addressed as one of the cross-cutting themes and within three evaluations questions (under Relevance and Effectiveness), as well as in the subsection on ethical considerations which notes issues to attend to when disability related data was collected. The stakeholder map included multiple entities that work with people with disabilities (PWD) and evaluation participants included representation from the Zanzibar Department of Disability Affairs. The analysis highlights UNFPA's work in this area including the development of a disability database, and the results of this analysis are reflected in one of the conclusions and one of the recommendations.</p>		
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 <u>assessment level</u> followed by main <u>comments</u> . (use 'shading' function to give cells corresponding colour)	
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Unsatisfactory
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 although the section headings, particularly those within findings, could be made more distinct. The writing is generally clear and easy to understand but there are some noticeable grammatical and spelling errors, and not all text is of consistent font size. In addition, the document is text-heavy and does not use clear visual aids to break up and summarize the text. For example, the images are primarily screen shots which are difficult to read and have not been cropped to remove unnecessary information. Rights-based language is not always used; for example it is generally not considered respectful to use the words 'crippled' or 'mentally retarded'.	
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)	No	The report, including the Executive Summary, is approximately 78 pages which exceeds the maximum length for CPEs of 70 pages.	
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)?	Partial	Annexes include the ToRs, a bibliography, an evaluation matrix and the data collection tools; missing is a complete list of interviewees. Although there is a stakeholder map included, it is not clear whether all of these stakeholders were interviewed.	
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	All of the core components for an executive summary are included; however the description of the methodology is general and does not provide much detail on the number and types of stakeholders engaged, nor on whether data was collected remotely or in-person. There are separate sections for findings and conclusions, though they are somewhat repetitive and do not provide additional beneficial analysis for this standalone section.	
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The executive summary is 6 pages long, slightly exceeding the maximum length. It is also text heavy in some places with several long paragraphs, one of which is more than one page long. Although the writing is clear, a more concise presentation would make this section more accessible.	
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 development context is clearly explained in Chapter 2, including constraints to progress on some indicators.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The evaluation discusses the intervention logic in Chapter 3, but does not assess it. The chapter on effectiveness briefly assesses the logic for each component (SRHR, Population Development, etc.).
<i>To ensure a rigorous design and methodology</i>		
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?	Yes	The evaluation matrix is briefly described in the text and provided as an annex, including evaluation questions by criterion, assumptions, indicators and appropriate data sources, and methods for data collection aligned with the indicators. The key findings are also detailed for each criterion. A further aspect of good practice is that the evaluators noted the changes made to the evaluation questions from what was proposed in the ToR.
4. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection are described and include key informant interviews, focus group discussions, and document review, and the justification for their selection is noted to be based on the indicators within the evaluation matrix.
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)?	Partial	A detailed stakeholder map is provided in Annex 4, which organizes institutional stakeholders by CPD output. The stakeholder groups consulted are also detailed in Table 2, however end beneficiaries, such as reproductive age girls, are not included despite the report stating that focus groups were conducted with this target group amongst other end beneficiaries. In addition, there is no mention of whether key stakeholders were consulted on recommendations [though it is stated that they were consulted to validate preliminary findings (p8) and that they 'would' be consulted in subsequent feedback rounds with the evaluation reference group (p69)].
6. Are the methods for analysis clearly described for all types of data?	Yes	The evaluation report notes that narrative analysis and contribution analysis were the primary methods of qualitative data analysis and that descriptive statistical methods were used for quantitative data. Each is briefly explained.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Partial	The methodological limitations are noted and what was done to minimize the limitations briefly described. The mitigation measures could be further elaborated on to better understand how they were addressing gaps and limitations. For example, the remote nature of interviews was mentioned as a limitation to quality data with only the use of technology noted as a mitigating factor. It also states that it was challenging to identify direct beneficiaries, however there is no information on the number of indirect beneficiaries engaged or the constraints or risks in how focus groups were conducted (it is briefly noted under section 1.3.2 that they were done in-person, but simultaneously noted to be done online). As such, it is not made clear whether focus groups were carried out in actuality or whether interviews with implementing partners were conducted instead as a mitigating measure.
8. Is the sampling strategy described?	Partial	The sampling strategy was purposive and described in section 1.3.5. However, as noted, the resulting sample is not included in the methodology section in terms of the numbers engaged or the location of beneficiary communities; without this information, the representativeness of the sample cannot be assessed and this is a major shortcoming of the report. It is stated that "cognizance was given to location in the choice of beneficiary communities to be selected for the evaluation" and the "variety of interventions supported by UNFPA Tanzania" was considered. Later in the relevance section, it does note the locations of focus groups (two in Mainland Tanzania, and two in Zanzibar).
9. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The methodology should enable the collection of disaggregated data through extensive document review as well as broad stakeholder consultation through interviews and focus groups with different population groups, however stakeholder lists are not disaggregated by sex or disability status. The context chapter does present sex-disaggregated data, and the findings also indicate that sex-disaggregated data was collected through the inclusion of quotes from both males and females.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	The methodology is appropriate for assessing cross-cutting issues based on the information available, which describes plans for focus groups with four different beneficiary groups, the use of rights-based approaches, and the consideration of the specific vulnerabilities of GBV survivors and persons with disability. Evaluation questions also assess whether the needs of marginalized and vulnerable groups were addressed and considered, and whether there were differential results across groups.
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 evaluation findings triangulate data across sources / methods, and the evaluation matrix shows clear intent to collect triangulated data for each evaluation question.

2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	Data sources are consistently identified and the reliability of sources is briefly discussed in the methodology section. However, as noted, the number of sources consulted for each method is not clearly mentioned.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	There is a subsection on Ethical Considerations which discusses the UNFPA and UNEG guidance that was followed, including the UNEG Ethical Guidelines and Code of Conduct. There is also a brief explanation of additional considerations made in conducting interviews where data on GBV and disability issues was collected. The data collection tools are included in Annex 5, and the introductory paragraphs make note of the ethical considerations followed. Data sources are not purposely identified by name or position, though they are identified by data collection method, location, and sex of respondent. However, some of the quotes in the report include reference the specific organization of the person being interviewed and/or the various roles they serve. It is recommended to only include quotes that do not clearly identify one person, unless the consent of that person to be identifiable was explicitly obtained.
4. Analysis and Findings	Yes No Partial	Assessment Level: Good
To ensure sound analysis and credible findings		
1. Are the findings substantiated by evidence?	Partial	The findings are typically substantiated by multiple lines of secondary/documentary and primary-level evidence, as well as both quantitative and qualitative data. However, in some cases, the challenges and lessons presented under effectiveness and under the 'leaving no one behind' section, while reasonable, are not substantiated with sufficient levels evidence (cited interviews, documents, etc.).
2. Is the basis for interpretations carefully described?	Yes	Throughout most of this section findings are clearly stated and then supported by explanatory text. The sub-section on sustainability provides a concise illustration of where a useful explanation is given for the finding, in this case on the relationship with the Government of Tanzania.
3. Is the analysis presented against the evaluation questions?	Partial	The evaluation questions are listed for each criteria, however they are not used to structure the analysis, which would have been useful for the report user to better understand the extent to which evaluation questions were responded to.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The methodology notes the limits to drawing cause-and-effect links between the intervention and end results due to the evaluation design, however evaluators do effectively use the information available to show what was expected to be done, what was done in practice, and extent of achievement of planned results (both quantitative and qualitative). The methodology section indicates that contribution analysis was used to assess both intended and unintended outcomes, and to some extent the latter was evident in the organization's response to the pandemic.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis shows differential outcomes for the various target groups. Where available, sex-disaggregated statistics are presented (for example, page 42 discussion of HIV infection rates).
6. Is the analysis presented against contextual factors?	Yes	The findings are clearly shown against contextual factors, frequently through the use text boxes to present the summary of challenges to implementation and achievement of indicators.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The analysis elaborates on cross-cutting issues of vulnerability, disability inclusion, gender equality and human rights, particularly under relevance and the discussion of challenges in the different intervention areas.
5. Conclusions	Yes No Partial	Assessment Level: Good
To assess the validity of conclusions		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions extend directly from the findings and show the question(s) and criteria on which each was based.
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	The conclusions adequately address cross-cutting issues related to gender and vulnerability based on other factors (youth, for example). However, they are numerous (17) and do not provide much additional analysis beyond the findings. Rather, they simply reiterate the bolded key findings statement provided there. This section could be more impactful if it presented a more concise set of higher-level conclusions that reflected the strengths and challenges for each criteria.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no apparent biased judgment from evaluators given the conclusions' direct link to findings.
6. Recommendations	Yes No Partial	Assessment Level: Very good
To ensure the usefulness and clarity of recommendations		

1. Do recommendations flow logically from conclusions?	Partial	The recommendations are directly linked to specific conclusions through numbering, however sometimes the logic behind the connection is not clear. For example, recommendation 13 on offering integrated interventions for women, youth and adolescents with a strong focus on vocational skills training is linked to conclusion 9 which speaks about issues related to sexual and reproductive health, including access to obstetric and neonatal care services.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations identify the intended users, including information on the technical and financial implications of 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?	Yes	The recommendations appear balanced and impartial.
4. Are the recommendations prioritized?	Partial	The recommendations are all prioritized, but the priority level is 'high' for all of them. Although the timeframe for implementation is also noted, which provides more variation on those of top priority based on the immediacy of the timeframe suggested, clearer prioritization would make the large number of recommendations (16) more manageable.
7. Gender	0 1 2 3 (**)	Assessment Level: Very 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?		<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) This was not evident in the objectives but the thematic scope directly addresses human rights and gender. (2)</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 and human rights was mainstreamed into other criteria through the evaluation questions. (3)</p> <p>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) There are dedicated evaluation questions and assumptions which address 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 assesses the results indicators, but does not clearly assess whether sufficient information was collected on progress on human rights and gender equality results. (1)</p>
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3) The methodology notes how gender and inclusion were addressed, and specifically how interviews considered ethical safeguards when collecting information on GBV and disability. However, the content detailed is more aligned with the type of description in an inception report, and it was not noted whether any interviews were conducted with survivors in actuality and whether safeguarding measures had to be employed. (2)</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) The evaluation methodology used mixed methods appropriate for evaluating GEEW considerations, however the sample size was not referenced to assess whether it was appropriate (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 and triangulation methods were applied. (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 reached a diverse group of stakeholders (girls of reproductive age, refugees, men who were part of action groups and other beneficiaries), though as noted above, the numbers reached aren't made clear. (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 and considered. (3)</p>
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		<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) A comprehensive overview of the specific social groups affected by various issues and the related normative instruments to address human rights and gender equality is provided. (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 include the voices of different social groups. (3)</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Unanticipated effects of the interventions on gender equality are addressed in respect to UNFPA being able to adapt emerging needs and priorities around the pandemic. (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) There are several recommendations oriented towards addressing equity and inclusion. (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	31	62		7
Overall assessment level of evaluation report	Good			
<p>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.</p> <p>The cross-cutting themes, including disability inclusion, are adequately addressed.</p>				
<p>Consideration of significant constraints (e.g. COVID-19 or civil unrest)</p> <p>The quality of this evaluation report has been hampered by exceptionally difficult circumstances: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p> <div style="background-color: #f0f0f0; height: 100px; margin-top: 5px;"></div>				