

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

Title of evaluation report: **EVALUATION DU 8ème PROGRAMME PAYS UNFPA CENTRAFRIQUE 2018-2022**

Overall quality of report: Good Date of assessment: **6 June 2022**

Overall comments: This evaluation provides a comprehensive assessment of UNFPA's 8th Country Programme in CAR. It is notable for its extensive input from stakeholders as over 200 people participated in the evaluation process, including rights holders in multiple locations, as well as for the careful assessment and suggested reformulation of the CP's theory of change and results framework. The reformulation included the development of proxy indicators in collaboration with the Evaluation Reference Group in order to better assess progress towards intended outcomes. The evaluators highlight using a theory of change approach to conduct contribution analysis and a participatory approach to engage a range of stakeholders in the evaluation process. The annexes include a comprehensive evaluation matrix with very detailed findings and data sources which enables the main report to be within the required page length, however it would have been helpful if there was clearer reference in the main report to the annexed information. Although the methodological approach was mostly solid, the lack of a stakeholder map, which is instead shown as a list of stakeholders met, is a shortcoming, as is the very general description of rights holders consulted. The findings and conclusions are clearly articulated and provide insight into the challenges and shortcomings of the CP but could better highlight its achievements including the covid response as well as unexpected outcomes. The recommendations appear useful for the formulation of the next CP. It is not clear if they had been validated by the Evaluation Reference Group which would help ensure that they were actionable. The evaluation was gender responsive with issues of gender equality and GBV being reflected throughout the report. Disability responsiveness is fully reflected in the evaluation questions (being integrated into EQ1, 3 and 5) and in the associated findings. It appears that there was intent to engage people with disabilities (PWD) in the evaluation process although it is not clear this was done. Within the conclusions and recommendations, disability issues are not considered as a distinct issue but rather are mainstreamed into the broader theme of vulnerability and marginalized groups.

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:	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)?	Partial	The report follows a logical structure. Each section is well defined. However, in some places the report is text heavy (i.e., the findings section of the Executive Summary, the findings summary text box on p 31, and the continuous text on p 33) and would be more accessible if there were additional subheadings, bolding, spacing between paragraphs, and inclusion of participant quotes. Although right-based language is used (referencing 'duty bearers' and 'rights holders'), so are terms such as 'deaf-mutes' which are not considered to be respectful.
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 65 pages long, 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 are presented in a separate volume. They contain the required elements plus the ToC for the overall programme and each programme area.

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 required elements are included, however it is good practice for the methodology to also indicate the number of evaluation participants and the stakeholder groups represented. This would be useful for understanding the scope and credibility of the evaluation process. It would also be helpful if the findings and conclusions were reformulated so that the evidence was presented in the findings and then the conclusions were overall statements. Alternatively, these two sections could be combined. As they are, there is some repetition between the two sections.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	This section is concisely presented and is just under 5 pages in length.

2. Design and Methodology	Yes No Partial	Assessment Level:	Good
<i>To ensure that the evaluation is put within its context</i>			
1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The country context and UNFPA CP response sections clearly articulate the challenges and constraints.	
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The ToC and results framework are presented and assessed in detail in Chapter 3. The intervention logic is also assessed as part of the analysis of EQ7 which looks at the institutional M&E system. The findings in the evaluation matrix state that the indicators were often not measurable and that the evaluators remedied this by working with the evaluation manager and ERG to identify proxy indicators (Vol 2 p 72).	
<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?	Partial	The evaluation approach is explained as being theory-based and participatory. The annexes include the complete evaluation matrix with very detailed and carefully presented findings that spans over 100 pages. The matrix has visual aids (including photos and tables), as well as a section on recommendations for each question. The matrix is only briefly referenced in the main report. As the matrix includes detailed findings and sources of evidence, it would have been helpful if the reader had been more clearly directed to this as a source of further information.	
4. Are the tools for data collection described and their choice justified?	Partial	The tools for interviews, FGDs and observations are briefly described and justified. However, the annexed KII and FGD guides each have an extensive number of questions and there is no indication of which will be directed to different stakeholder groups.	
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)?	No	The stakeholder groups participating in the evaluation are identified as "representatives of government, civil society organizations, implementing partners, the private sector, academia, other UN agencies, donors and, most importantly, rights holders (women, adolescents and young people)" in the discussion on participatory approach in section 1.3.2. However a stakeholder map identifying the range of groups involved in the CP is not included even though there is mention of sampling being done based on stakeholder mapping. If this was provided as part of the evaluation report it would be feasible to assess the representativeness of the final sample. Evaluators note that an ERG was involved in the finalization of evaluation questions but no further explanation of either the make up of the ERG or their further participation in the evaluation process is provided.	
6. Are the methods for analysis clearly described for all types of data?	Partial	Section 1.3.2 is entitled Data Collection and Analysis Methods. It addresses the preliminary analysis and in-depth analysis phases but the description is at a general level. It is not clear whether a systematic process, such as coding, or software were utilized.	
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	Limitations (including security, data availability and quality, timing of field mission and impact of covid) and mitigation strategies are identified.	
8. Is the sampling strategy described?	Yes	A purposive sample was used with the evaluators noting that the process followed the guidance in the UNFPA manual. The selection criteria are listed. The resultant sample of locations for field visits is listed, along with the number of people consulted by either FGDs or KIIs, in table 2. Table 2 disaggregates the total of 218 participants by stakeholder group.	
9. Does the methodology enable the collection and analysis of disaggregated data?	No	The approach should have enabled disaggregated data to be collected but it is not clear this was done. The annexed FGD guide directs facilitators to identify the group being engaged and lists options that suggests that groupings were by age (10-24, 20-24, 25-49). There is also a category for 'vulnerable people (15-24 years old)' but no further explanation of what groups are considered vulnerable (such as people with disabilities) is given. The total number of evaluation participants is not gender disaggregated, and there is no indication of how many FGD participants fell into each of the other categories.	
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 country context section presents data disaggregated by gender, age, and refugee status. In terms of the primary data collected, the evaluation questions, the diverse range of stakeholders consulted, and the reported findings suggest that the approach was appropriate to assess cross cutting issues, including disability.	

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 matrix shows the intent to collect multiple sources of data for each question, and the findings sections in the matrix provides evidence of triangulation.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	Document sources are described and appear appropriate. The list of stakeholders participating in KIs and FGDs also appear appropriately diverse and reliable, although it would be helpful if more information was provided on participating rights holders. As noted in 2.9, the FGD guide indicates that a range of beneficiaries were to be involved, but there is no explanation of who was eventually engaged beyond the total number of rights holders.	
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	There is a discussion on ethical consideration that addresses issues such as confidentiality, anonymity and freedom to not answer all questions. Adherence to UNEG ethical principles is also mentioned. However, the annexed interview and FGD guides do not refer to ethical considerations. It is also problematic that the names and contact information for young people engaged in FGDs is included in the annexed list of evaluation participants.	
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	Findings are well sourced in the evaluation matrix, and there is some citing of document sources in the main report.	
2. Is the basis for interpretations carefully described?	Partial	There is a concise but solid explanation of achievements and of areas where intended performance fell short. The evaluators do particularly well in explaining cases where indicators needed to be reformulated. However, in most cases, the summarized findings in the main report only capture the shortcomings and not the accomplishments.	
3. Is the analysis presented against the evaluation questions?	Yes	The findings are organized by evaluation question. A summary of the findings is presented in a text box.	
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Partial	EQ4 looks at the extent to which outputs have contributed to outcomes (to the extent that the intended results were measurable). Tables showing performance against indicators (both the original and proxy indicators developed by the evaluation team) are also provided. Although the ToR requests that unintended results be examined, the report does not explicitly address these.	
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	This is done in several places. For example, the discussion on output 3 under Effectiveness notes that the CP had difficulty in reaching adolescents and young people in IDP camps and hot spots in the country.	
6. Is the analysis presented against contextual factors?	Yes	The context is consistently provided, particularly in terms of the office's covid response and how budgeting and activities were adapted, as well as the capacity-related challenges for partners to implement and sustain activities.	
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Cross-cutting issues are discussed throughout the findings. An example is the case study presented in EQ10 UNFPA's humanitarian response on the integrated GBV medical management services at Bossembele hospital.	
5. Conclusions			
	Yes No Partial	Assessment Level:	Fair
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Partial	Each conclusion is linked to the respective evaluation question(s). However, conclusions do not reflect the findings for EQ 3 which asked about the covid response (this was also specified as part of the scope of the evaluation). Conclusions could also have been more balanced by giving more emphasis to the accomplishments of the CP (such as the achievements in family planning and success in reaching people with disabilities).	
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	This section mostly summarizes the findings, particularly those related to the challenges and shortcomings of the CP. It is good practice for conclusions to also be forward-looking and show opportunities to build upon (for example on areas where there were achievement or over-achievement on indicators). Cross-cutting issues are addressed, with the exception of disability inclusion.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of evaluator bias.	

6. Recommendations	Yes No Partial	Assessment Level:	Good
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The recommendations specify the conclusions on which they are based.	
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	The intended users/implementors are clearly identified for each recommendation. In respect to actionability, most include sub-recommendations to guide implementation. However, there is some contradiction in respect to financial implications. Recommendation 3 suggests limiting the number of geographic areas to be covered due to limited resources, however other recommendations such as #4 and #10 suggest there should be substantially larger investments to meet the growing needs of populations.	
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	The recommendations appear balanced in terms of immediate through to long term timeframes. They explicitly address cross-cutting issues of vulnerability and gender equality, but there is not specific reference to human rights and disability inclusion.	
4. Are the recommendations prioritized?	Yes	Priority levels of 1 and 2 are assigned, with 7 of the 11 recommendations being priority 1.	
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?		<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, including GE, is part of the scope. = 3</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) GE is fully mainstreamed into the criteria. = 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) Five of the 10 main questions assess various aspects of how GEEW was integrated into the CP. = 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) EQ7 looks at the M&E systems and the findings in the evaluation indicating the system did not adequately enable such data to be collected. = 3</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) Evaluators note that the process respected human rights and GE but could have been more specific about how this was done. A notable shortcoming is that evaluation participants were not disaggregated by gender. = 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) Mixed methods were used and the sample size of rights holders (approximately 50% of all participants) appears appropriate. = 3</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) The approach involved multiple methods of data collection and multiple sources. Triangulation was clearly evident. = 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) Participants included a suitable range of rights holders and duty bearers from six of the seven prefectures where UNFPA works. However, rights holders could be more clearly described. = 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) Adherence to UNEG and other ethical guidelines are mentioned and practices such as informing participants of confidentiality and anonymity are identified. However, this section could have gone further by discussing special consideration in undertaking FGDs to ensure participants felt comfortable and safe. It is also problematic that the list of stakeholders consulted includes the names and phone numbers of youth beneficiaries. = 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 adequately covers issues faced by women and girls. = 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 perspectives of different groups, including rights holders, are evident in the evaluation matrix but could be better reflected in the main report. In addition the findings do not show disaggregated quantitative data by gender. = 2</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Unanticipated effects are not explicitly addressed. = 0</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) Several recommendations suggests ways to further work on GEEW and GBV. = 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	7	82	11	

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.
 These issues were all explicitly included in the evaluation scope. As noted above, disability was included in the evaluation questions and analysis but could have been better highlighted in conclusions and recommendations.

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: