



Organizational unit:   Year of report: 2021

Title of evaluation report: UNFPA Afghanistan 4th Country Programme 2015 - 2021 Final Evaluation Report

Overall quality of report: Very good Date of assessment: 8 June 2021

**Overall comments:** The evaluation provides a thorough examination of the Country Programme. The findings are well substantiated and give a very detailed account of the programme activities, but perhaps more than is necessary for an evaluation as the report ends up being quite lengthy. The Conclusions are particularly well formulated and help to synthesize the main results. The set of recommendations appears useful for informing the development of the next CP. The evaluation is notable for practices followed to enable a level of in-person data collection during covid. These included using local consultants who were able to visit programme participants at their residences when gatherings were restricted, and using masks and hand sanitizer to minimize exposure risks during interviews. The evaluation also did well in examining the extent to which the CP was disability inclusive, with disability-related issues being evident in the findings, conclusions and recommendations. There are several areas where the evaluation could be improved. In respect to Design and Methodology, it is not clear how the intent of the evaluation to include the participation of vulnerable groups was achieved - the sampling framework did not indicate numbers of rightsholders included and their voices were not evident in the data presented. The gender breakdown of evaluation participants was also not provided. The overall presentation of the report could have been improved by using a clearer and larger-sized font, by having more concise and less text-dense findings, and by final editing.

**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)*

<b>1. Structure and Clarity of Reporting</b>	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 structure is logical, however the presentation could be improved. Very small font is used and there are long sections of uninterrupted text, both of which make the report more challenging to read. The report would have benefitted from final editing as there are a number of spelling and grammatical errors.
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 is 74 pages, excluding the Executive Summary and Annexes. If a regular font size were used it would be considerably longer.
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 plus a stakeholder map and the Atlas project listing - for a total of 7 annexes.

*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?	Yes	The summary serves as a stand-alone section. It includes all of the required elements. There is a minor editing issue where the evaluation is referred to as the CP rather than the CPE in the description of Purpose.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	At 4.5 pages, it is within the limits.

<b>2. Design and Methodology</b>	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	There is a solid chapter on the country context and UNFPA's response.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The evaluators describe how a theory-based (and participatory) approach was used to guide the evaluation, and how they assessed and adjusted the ToC. Changes were made to the depiction of causal linkages, and modes of engagement and strategies for each thematic area were added.

*To ensure a rigorous design and methodology*

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 framework and its use in guiding the evaluation process is described in the main report. The attached matrix incorporates all of the required elements.
4. Are the tools for data collection described and their choice justified?	Yes	Data tools and the reasons for their selection are described.

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 mapping process was included. Stakeholder consultation is covered the methodology section and it involved the participation of the ERG in several stages of the evaluation process including 'consultation' on recommendations. What is less clear is the engagement of rightsholders in the evaluation process. In the Executive Summary it is noted that direct and indirect beneficiaries were consulted, and in the methodology section on Evaluation Approach it is noted that "particular attention was paid to ensuring participation of women, adolescent girls and young people, especially from vulnerable and marginalized communities". However, the total number of participants by stakeholder group is not provided in the main text and in the annexed list of participants there are approximately 40 beneficiaries listed but other than a few noted as being members of the Youth Parliament, the remainder are healthcare workers. Therefore, it is not clear vulnerable groups were actually represented and the list is not gender-disaggregated to verify the extent that the stated gender representation was achieved.
6. Are the methods for analysis clearly described for all types of data?	Partial	There is a general explanation given for how the different types of data were analyzed. For example, contribution and thematic content analysis was used for interview data, contribution analysis based on the ToC was to be used for documentary evidence, and descriptive statistics for quantitative data. However, more details of each would have been useful and there is no evidence of presentation of the latter in the findings.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	Five limitations, including those related to Covid, and their mitigation strategies are described.
8. Is the sampling strategy described?	Partial	The sampling strategy (purposive) is well described for the FGDs and KIs. The universe is clear, as is the number of individual and group interviews and FGDs. However, the number of people in each stakeholder group is not given in the main text (although can be calculated from the annexed list of participants). It is noted that site visits were determined in consultation with "the team on the ground" and that site visits were restricted to Kabul but although two examples of selected interventions for visits are provided (youth activities or safe GBV houses), no further description of the sites is given.
9. Does the methodology enable the collection and analysis of disaggregated data?	No	A listed limitation is the availability of adequately disaggregated programme data. However, it is not clear what steps were taken to ensure that evaluation data was collected in this way. The evaluation matrix does not include indicators calling for disaggregated data, and the tools do not require this. Just two instances disaggregated data were seen (one on p 34) and these were from annual reporting.
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 design is generally appropriate for such an assessment. In addition to a specific question on the extent that a gender-responsive and HRBA was used in programming, there is an indicator on the extent to which targeted vulnerable populations, including people with disabilities, were consulted in the CP design and activities. The design also takes into account the humanitarian setting by including the criteria of coverage and connectedness.
<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	This is done. Multiple sources are cited; these are frequently each identified in brackets right after the respective finding.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	Qualitative data are attributed to specific stakeholder groups but there is minimal citing of specific document sources, and instead the source is in most cases just noted generally as 'document review'. Being more clear about the document source would be useful given that one of the methodological limitations identified is in regards to the quality and amount of relevant documents and reports, and the potential for bias within secondary data produced by the CO.

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 guidance is stated. However the evaluators also explain specific considerations including use of local consultants to address language barriers, having a female collect data from female participants, social distancing, ensuring informed consent, etc.
<b>4. Analysis and Findings</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	The findings are very descriptive and are backed up by ample evidence.
2. Is the basis for interpretations carefully described?	Yes	This is done. For example, the relevance of the CP to country context is backed up by citing activities that aligned with cultural practices (i.e., referring to family planning as birth spacing)
3. Is the analysis presented against the evaluation questions?	Yes	The evaluation questions and a summary of findings appear before the relevant analysis for each.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluators describe in detail the intervention activities that supported the findings. One unintended result noted was that a manual developed for the police on VAWG has been used more widely than initially intended (now used in the Police Training Academy). There is also a subsection on Unintended Effects that covers this for each of the four programme areas.
5. Does the analysis show different outcomes for different target groups, as relevant?	No	The analysis describes programming directed towards different groups but does not report on results.
6. Is the analysis presented against contextual factors?	Yes	Contextual factors are frequently identified such as the challenges hampering effective response to emergencies, and the implications of the significant cuts in US funding. The text also describes how UNFPA has adapted its activities to overcome challenges (such as hiring male relatives to work alongside needed women healthcare workers).
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Section 4.3.5 looks specifically at the integration of gender and HRBA in the CP. This includes information on the programme's reach to vulnerable groups and a reasonably in-depth look at how disability inclusion was addressed (i.e., an indicator in the socio-demographic and economic survey (SDES), efforts to include people with disabilities within the Youth Parliament programme.)
<b>5. Conclusions</b>		
	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 linkage is evident and is confirmed by a notation of the relevant evaluation question numbers informing each conclusion.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Yes	The conclusions do provide a higher-level view of the programme, particularly of the areas where there were challenges/shortcomings. Cross-cutting issues are addressed.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no indication of bias.
<b>6. Recommendations</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	The flow is logical and the relevant conclusion numbers are given for each recommendation.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	These are clearly targeted to UNFPA for consideration in the development of the next CP. Operational Implications are clearly specified for each.
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	These appear balanced, impartial and address the cross-cutting issues including people with disabilities.
4. Are the recommendations prioritized?	Yes	Priority levels are assigned to each with most being High priority.

<b>7. Gender</b>	0 1 2 3 (**)	Assessment Level:	<b>Very good</b>
------------------	-----------------------	-------------------	------------------

To assess the integration of Gender Equality and Empowerment of Women (GEEW) (\*)

<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>	2	<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> The intent to ensure the evaluation assesses HRGE is not mentioned in the scope or objectives. = 0  <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> Gender mainstreaming is noted as being a cross-cutting theme = 3  <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 specific question on this - EQ4 on Effectiveness. = 3  <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> Although this topic is not featured in the evaluation matrix, the evaluators do make this assessment. = 2</p>
<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	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> The relevant guidance is noted as being used and it is stated that efforts were made to ensure adequate representation of women and men. However, very little gender disaggregated data is collected, including of the breakdown of evaluation participants. = 1  <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> Mixed methods were used, although most primary data was qualitative. Specific sample size is not provided in the main text = 2  <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> Data was collected from a range of sources and triangulation was evident. = 3  <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> Different stakeholder groups were consulted. The sampling frame also supposedly included the most vulnerable but this could not be confirmed = 1  <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 were noted, including confidentiality - however, the names of participants were listed in the annex = 1</p>
<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	3	<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> This is well done. = 3  <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 evaluation falls short here - although in a few cases beneficiaries are noted as being one of the data sources, their perspectives are not highlighted. Quotes would be useful. = 1  <b>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)</b> This is done. = 3  <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> This is done. = 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>82</b>	<b>18</b>		
<b>Overall assessment level of evaluation report</b>	<b>Very good</b>			