



Organizational unit: **Albania Country Office** Year of report: **2021**

Title of evaluation report: **Evaluation of the 4th UNFPA Country Programme for Albania (2017-2021)**

Overall quality of report: **Good** Date of assessment: **20 April 2021**

**Overall comments:** This is a comprehensive independent evaluation of the 4th UNFPA Country Programme for Albania with the programme and each component being assessed in terms of its relevance, effectiveness, efficiency, sustainability, coordination and added value. The evaluation uses a mix-methods approach which is appropriate for assessing the four components: (1) Sexual and Reproduction Health, (2) Adolescent and Youth, (3) Gender, and (4) Population and Development. The findings are very well supported by both qualitative and quantitative data. In fact, notable strengths of the report include the meticulous citing of data sources in the findings and the presentation of key findings in the evaluation matrix. Both the conclusions and recommendations clearly flow from the findings section.

However, there are several areas where the report could more closely align with UNFPA's standards for high quality (very good) evaluations. These include: collapsing the number of conclusions (25) in to a more manageable number; providing an explanation for why the case study requested in the ToR was not included in the design; being more explicit about how the evaluation process was gender responsive; and, given the emphasis placed on a "broadened evidence-base for the design of the next programming cycle", amplifying the "lessons learned". It is briefly noted that the CP addressed issues of disability, however there is no further analysis provided. In addition, although the report is generally easy to read and to understand, it would benefit from final editing and the inclusion of more space between sections and some of the text.

**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:	<b>Fair</b>
----------------------	--	-------------------	-------------

*To ensure the report is comprehensive and user-friendly*

<b>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)?</b>	Partial	Although the report is easy to navigate and understand, there are a number of grammatical and spelling errors, including in the Executive Summary. The distinction between the different sections of the report could be more apparent if the chapter headings had a larger font size or color. Although there is good inclusion of visual aids, two of the graphics are too small to easily read, including one on SDGs. The report is also quite text heavy in places and includes a very detailed account of the evaluation process in the methodology section that appears to have been taken from the inception report as it refers to future processes.
<b>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)</b>	Yes	The report is 66 pages, including the executive summary, which meets the preferred length for UNFPA CPE reports. [However, it is noted that the ToR states the report should be "a maximum of 50 pages" (plus up to 70 pages for "each case study"). To help meet this requirement, some of the detailed text could have been moved to the Annexes.]
<b>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)?</b>	No	The annexes contain most of the specific items, with the exception of the data collection tools used.

*Executive summary*

<b>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?</b>	Yes	The executive summary is written as a stand-alone section, and presents all required elements.
<b>5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	Partial	The executive summary is 6 pages, one page more than UNFPA's maximum length. However, the text is quite dense and there is little spacing. Readability would be increased if it were more concisely written - one way of doing this would be to provide less detailed findings and conclusions.

**2. Design and Methodology**

Yes No Partial		Assessment Level:	<b>Very good</b>
----------------------	--	-------------------	------------------

*To ensure that the evaluation is put within its context*

<b>1. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	There is a lengthy section that clearly describes the country's development and institutional context of the evaluation, and the constraints involved.
<b>2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?</b>	Yes	The CP programme intervention logic (theory of change) is included in the chapter on UNFPA's response with the programme ToCs for two of four components appearing in Annex 7. It is a bit confusing, as the third paragraph of findings indicates that Annex 8 includes the ToCs for each programme areas. Analysis of the adequacy of the logic is done in several places under Effectiveness, with the issues raised being carried through to Conclusions and Recommendations.

*To ensure a rigorous design and methodology*

<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>	Yes	The evaluation framework is clearly described in the text, and the evaluation matrix is presented in Annex 4. The matrix includes each of the respective items. The matrix is not only used as a data collection tool but it is also used for analysis purposes, with the actual data for summarizing evaluation results being included in the matrix.
<b>4. Are the tools for data collection described and their choice justified?</b>	Yes	The tools for data collection involve four main methods - 1) desk review, 2) key informant semi-structured group and individual interviews, 3) online survey, and, due to COVID-19, 4) a limited number of visits to selected implementing partners. These data collection tools are described and the rationale for their selection is noted in the text.

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)?	Yes	There is a description of stakeholders in the main report and a stakeholder map is included in Annex 6. The process for selecting the sample of stakeholders is based on the mapping exercise. The report describes the overall stakeholder consultation as "an inclusive approach" and describes in great detail how the Evaluation Team regularly took into account comments from various key stakeholders on the different versions of the report during the review process, including the Evaluation Reference Group. Although not specified, it can be assumed that the draft reports included recommendations.
6. Are the methods for analysis clearly described for all types of data?	Partial	There is a paragraph on data analysis that lists multiple types of analysis processes used - document analysis, ToC and contribution analysis, content analysis, descriptive statistics. Although this appears to be extensive, it would be helpful if there was reference to the mechanics of these processes, including linking the type of data to the analysis process used.
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 evaluation report has a brief section on Evaluation Limitations that notes the restrictions of the pandemic. Mitigation efforts are touched on. What is not clear is how the case study(ies) requested in the ToR were treated or if they were done at all. It is noted in the Methodology section that it was not feasible to visit municipalities where the Fund works due to limited time and resources and the COVID-19 pandemic, but it is not clear if those were the reasons for not doing the case study(ies).
8. Is the sampling strategy described?	Yes	The methodological approach to determining the sample of stakeholders consulted and sites visited is thoroughly described. An explanation of the criteria for selecting the stakeholders is given and notes that both purposeful sampling and convenience sampling were applied. The resulting sample is reported.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	Evaluation participants are disaggregated by programme area and gender for each of the primary data collection processes. Data is disaggregated to some extent in the contextual analysis and comparative analysis.
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 mixed methods approach used is appropriate for assessing cross-cutting issues. Evaluation questions cover HRBA and perspectives are collected from a range of stakeholder groups. Disability is mentioned several times.
<b>3. Reliability of Data</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Specific mention is made of data triangulation of the multiple sources and methods of data collection. All of the findings used multiple sources and methods.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation clearly identifies and makes use of reliable qualitative and quantitative data sources. The evaluators discuss the reliability, or lack thereof, of both when it is an issue. For example, the evaluation notes that there is not full congruency between data sources when such is the case, or identifies gaps and challenges in the data when appropriate.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The evaluation notes that it closely adhered to the UNEG Conduct and Ethical Guidelines for Evaluation and briefly explains how the evaluation team endeavored to take these considerations into account during the conduct of the evaluation.
<b>4. Analysis and Findings</b>		
	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	The evaluation findings are substantiated by evidence in the text for each finding. The evaluators were meticulous in citing sources with even the respective KII number of survey question being identified in the footnotes.
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretations are shown with the widespread use of such terms as "according to respondents", "interviewed respondents stated", "on-line survey respondents indicated", "evidence from desk review showed", and "document analysis and key informant interviews showed".
3. Is the analysis presented against the evaluation questions?	Yes	The findings are presented against each of the evaluation questions and these are organized according to the respective evaluation criteria.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Partial	The cause-effect is clear in the findings, particularly in the effectiveness criteria - although the evaluation notes that "the UNFPA partners and target population groups do not necessarily attribute the interventions to the UNFPA and/or know about them". While not specifically citing unintended outcomes, the evaluation does make statements such as "at the same time this UNFPA approach has ...", the UNFPA approach "led to change of traditional socio-economic norms," and "UNFPA has contributed to three other major developments...".
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings analyze and show differences across relevant outcomes for different target groups.
6. Is the analysis presented against contextual factors?	Yes	The analysis and findings in the evaluation report are presented against contextual factors at the national, regional and local levels.
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 dealt with in the findings, conclusions and recommendations. Persons with disabilities are briefly mentioned as being addressed in the CP in the finding on EQ9.
<b>5. Conclusions</b>		
	Yes No Partial	Assessment Level: <b>Good</b>
<i>To assess the validity of conclusions</i>		

1. Do the conclusions flow clearly from the findings?	Yes	Conclusions are based on the analysis of the Findings and are treated at both the strategic and programmatic levels. Each conclusion is followed by the number of the appropriate evaluation question(s), the evaluation criteria, as well as the associated 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?	Partial	There are 24 conclusions. Although they are succinctly stated, good practice would be to just highlight the highest-level conclusions and fold others into these as feasible. As stated, they do convey the essence of the findings and cover cross-cutting issues of gender and vulnerability.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias.
<b>6. Recommendations</b>		
	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Partial	The recommendations logically flow from the conclusions. In addition to 6 programmatic recommendations, there are just three strategic recommendations even though one of the three specific objectives of the evaluation is "to draw key lessons from the past and current operation to provide a set of clear, specific and action-oriented forward looking strategic recommendations in light of agenda 2020 for the next programming cycle". This suggests there should have been either a more robust set of strategic recommendations in light of the importance given to "draw key lessons" or a specific set of "lessons learnt" as a separate section distinct from the Conclusion and Recommendation sections, with a notation explaining why they were not included in the strategic recommendations.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations are targeted to either the Country or the Regional Office or both. They are action oriented with operation implications 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?	Partial	The recommendations appear balanced and impartial. They do address cross-cutting issues such as gender equality, but not other issues such as equity and disability.
4. Are the recommendations prioritized?	Yes	The recommendations are prioritized as either "High" or "Medium."
<b>7. Gender</b>		
	0 1 2 3 (**)	Assessment Level: <b>Good</b>
<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><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> HRGE is not mentioned in the overall or specific objectives but is included as part of the scope of the evaluation. = 2</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> Gender equality is mainstreamed under Effectiveness and as a "cross-cutting lens". = 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> This is covered in EQ 3. = 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> This is taken up somewhat in the evaluation matrix. = 1</p>
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		<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> There is little clarity on the way this was done, other than the statement that GEEW was employed as a "cross-cutting lens". = 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> A mixed methods approach is used and the evaluation team states that it adopted "an inclusive approach involving a broad range of partners and stakeholders". = 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> A wide variety of sources and a range of processes were used to meet those criteria. = 3</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> Rights holders were included but it is difficult to determine if this included the most vulnerable. = 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> The evaluation team subscribed to the ethical standards and was "based on a person-centered approach" by reviewing cultural and religious sensitivities existing in Albania. = 3</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 country context section focuses on HRGE and social issues and includes a significant amount of gender and age-disaggregated data. = 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). This type of analysis is not done. The evaluation, however, does well in addressing how the needs of a range of social groups were met. = 2</p> <p>2 c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3). Only indirectly are the effects noted - e.g., the UNFPA approach led to change of traditional socio-cultural norms, such as engagement of youth in drafting a law on youth. = 2</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). A specific recommendation (to continue to support combating GBV) was included. = 2</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>24</b>	<b>69</b>	<b>7</b>	

Overall assessment level of evaluation report **Good**

The evaluation integrates adequately cross cutting issues of gender equality and human rights, even though is not included as part of the evaluation objective.

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

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

If yes, please explain:  
 Although planned earlier, the evaluation was carried out during the pandemic. The evaluators managed to adapt data collection and still consult a good range of stakeholders. It is less clear if the requested case study was not carried out because of the pandemic.