

Organizational unit:		Year of report:	2020
Title of evaluation report:	Evaluation of the First (2011-2015) and the Second (2016-2020) Country Programmes of UNFPA Belarus		
Overall quality of report:	Very Good	Date of assessment:	2020 November
Overall comments:	The evaluation is complex in covering two country programme periods over a ten year period. It is thorough in examining expected results that were to be obtained during the periods and shows which were achieved and which needed further work in the next cycle, the elaboration of which was one of the objectives of the evaluation. The methodology was reliant on extensive document review and a purposive sample of key informants, and utilized several focus groups with beneficiaries. No primary quantitative data was obtained. A detailed evaluation matrix set out evaluation questions and sub-questions around which the findings were organized. This permitted drawing a solid set of combined conclusions and recommendation which were practical, but not prioritized.		
Assessment Levels	Very Good	Good	Fair
	strong, above average, best practice	satisfactory, respectable	with some weaknesses, still acceptable
			Unsatisfactory
			weak, does not meet minimal quality standards

Quality Assessment Criteria	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>		
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Fair
<i>To ensure the report is comprehensive and user-friendly</i>			
1. 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?	Partial	The report is well presented and is generally well written with only a few editing omissions. However, it could be more accessible to wider audiences with more consistent use of shorter sentences.	
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 report is just under the maximum number of pages for CP evaluations.	
3. Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Yes	A logical structure is followed.	
4. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological tools used (e.g. interview guides; focus group notes, outline of surveys) as well as information on the stakeholder consultation process?	Partial	The annexes include the required components except for the data collection tools used.	
<i>Executive summary</i>			
5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?	Partial	The Executive Summary presents the context, audience, process and main results of the evaluation. However, the methodological scope could be more detailed by providing further information about primary data sources including number of people participating in KIIs and FGDs and stakeholder groups represented.	

6. Is there a clear structure of the executive summary, (i.e. i) Purpose; ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?	Yes	The required structure is followed.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The ES is 4 pages in length and concisely presented.
2. Design and Methodology		
	Yes No Partial	Assessment Level: Fair
<i>To ensure that the evaluation is put within its context</i>		
1. Does the evaluation describe the target audience for the evaluation?	Yes	The intended audience is described as being various UNFPA stakeholders, the different types of national partners, UNCT and donors.
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	These issues are well covered.
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes	The evaluators reconstructed ToCs for each programme area for both CP 1 and CP 2. These are comprehensive covering needs, inputs and activities through to impact and objectives.
<i>To ensure a rigorous design and methodology</i>		
4. 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 framework is addressed in the text. The annexed matrix is very detailed covering assumptions, indicative resources, indicators, sources or information, data collection methods and summarized findings for each evaluation question.
5. Are the tools for data collection described and their choice justified?	Yes	The evaluation used the typical data collection tools. These are described and are noted as being designed around the assumptions and indicators found in the evaluation matrix.
6. 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	The annex includes a stakeholder map that clearly shows the stakeholders involved for each output area and project. The methodology covers stakeholder involvement in the review and validation of findings. It is noted that the draft report was shared with UNFPA CO but there is no specific mention on consultation in the development of recommendations.
7. Are the methods for analysis clearly described for all types of data?	No	It is briefly noted that counterfactual analysis was applied where possible to explore cause and effect relationships, however there are no details on how this was done. There is no explanation for how each type of data was analyzed.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The main evaluation limitations and mitigation measures are shown in Table 3 and accompanying text. Bias is noted as being addressed through triangulation of sources under the subsection on Ethics.
9. Is the sampling strategy described?	Yes	This is well described. The sampling approach is noted as being based on the stakeholder map and following the guidance in the UNFPA Evaluation Handbook. The criteria and process for selecting stakeholders and locations to be covered is clearly outlined in Table 4 and accompanying text.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	This is not explicitly addressed in the methodology and the data collection tools are not attached to see how disaggregated data may have been collected. However there is some evidence this was done as the number of women and men participating in the FGDs is provided, and the Context section includes gender disaggregated data from document review. In the Findings, the perspectives from different stakeholder groups, including rights holders, of their needs and experiences are discussed separately.

1. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The evaluation effectively covers cross-cutting issues of gender, equity and vulnerability in the questions - including if the SRH services reach those "furthest behind". EQ1 covers the extent that rights holders were consulted as part of program design.
3. Reliability of Data		
	Yes No Partial	Assessment Level: Very good
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Multiple data sources are frequently cited as the basis for findings
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The sources of data are clear but primary data are mostly qualitative. There is an extensive and comprehensive list of documents and stakeholders consulted, including rights holders. One of the limitations notes that recall was an issue for collecting primary data on CP 1 and that therefore there was a greater reliance on document review. In the findings section, the evaluators make clear where data is limited to a specific population and cannot be extrapolated to the entire country (i.e. paragraph 2, page 49).
3. Did the evaluation make explicit any possible limitations (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues?	Yes	The discussion on limitations and mitigation addresses the limited availability of data particularly for the 1st CP.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The subsection on Ethics addresses confidentiality, avoidance of harm, sensitivity to local beliefs and customs, and other issues relevant to ethical practices for data collection. Although the evaluators do not go further by discussing limitations of the process, it is noted that English-Russian translation services were provided.
4. Analysis and Findings		
	Yes No Partial	Assessment Level: Very good
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	The basis for each of the findings is very clear.
2. Is the basis for interpretations carefully described?	Yes	The discussion and analysis for each question is thorough. Under Effectiveness, there are output indicator tables that show baseline, target and end line/milestone information.
3. Is the analysis presented against the evaluation questions?	Yes	The evaluation questions and indicators are highlighted for each criteria and are followed by the presentation and analysis of findings.
4. Is the analysis transparent about the sources and quality of data?	Yes	Data sources are clear. Footnotes are used throughout to cite source documents and the specific KIs from which data was obtained. The evaluators use bolding to highlight perspectives gained from key informants.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	EQ4 specifically addresses the extent to which outputs contributed to planned outcomes. This is done for each program area. Unintended results are addressed to some extent in discussions about program development - for example under improving national capacity for a multisector response to GBV and DV, it is noted that fines for perpetrators ended up negatively effectively household income (p 56).

6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	This is done. One example is in regards to UNFPA's contribution to the collection and open dissemination of data from surveys and other capacity building activities - a finding is that such data is not widely used outside of the science community and its potential is not well understood by other stakeholders. Another is the discussion on the work on the DV law, and the perspectives of different stakeholders, including church-based groups, on the path forwards (p. 57).
7. Is the analysis presented against contextual factors?	Yes	The context is consistently provided, including for example, the government and other partners' perspectives and contributions in the different program areas.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	This is done including in how UNFPA has directed more attention to issues of aging, and how the YFHS program could do more to reach key populations.
5. Conclusions		
	Yes No Partial	Assessment Level: Very good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions reflect the findings. They are organized by three categories: strategic level, programmatic level, and operations and management. The relevant criteria for each category is listed which assists in making the linkages clear.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Yes	The conclusions are framed in a way that shows the overall strengths and areas for improvement for the CP.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of 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 are also organized by the categories presented in the conclusions. The relevant evaluation questions are also highlighted.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The intended users are made clear and operational considerations/recommendations are detailed for each.
3. Do recommendations appear balanced and impartial?	Yes	This criteria is met.
4. Is a timeframe for implementation proposed?	Yes	It is noted at the beginning of the Conclusions and Recommendations section that the learnings from the evaluation are to be considered in the planning the next programme cycle.
5. Are the recommendations prioritized?	Partial	All main recommendations appear important but they are not explicitly prioritized. Where prioritization would have been most relevant is in the ordering of the operational sub-recommendations which are quite numerous.

7. Gender	0 1 2 3 (**)	Assessment Level:	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?		2	<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 HRGE considerations is not explicitly stated in the objectives but it is noted as a cross-cutting issue. = 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) It is mainstreamed but the presentation is not clear on this. Gender is stated as being a separate criteria in the methodology discussion (p 18-19) but it is not included in the list of criteria covered in the Executive Summary. Since UNFPA's Gender Equality programme is part of the evaluation, there is a specific section under findings on this. = 2</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) GEEW is integrated into multiple questions. = 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 matrix includes indicators that look at national capacity to produce gender-disaggregated data but much of the data collected by the evaluation was not disaggregated (i.e. EQ 4.2 Output 6 looks at number of specialists who organize peer-to-peer training, number of state representatives sensitized, number of young people covered by SRH activities - but none of these indicators are require gender disaggregated data). = 1</p>
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		2	<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) There is not an explicit discussion on how the methodology was gender responsive. The total number of evaluation participants were not gender disaggregated, although this was done for the FGD participants. = 1</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) Overall, the methods used appear appropriate - a mixed methods approach was followed although the only quantitative data was from document sources. The sample size was adequate. = 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) A diverse range of sources were used = 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 sampling frame included rights holders from vulnerable populations. The finding included general perspectives obtained from FGDs but use of quotes or quantitative data from these discussions would be useful. = 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 are described. As noted above, one concern related to confidentiality is that mission reports used as evidence included the names of the staff who wrote them. = 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 Key Facts chart at the beginning of the report includes relevant gender statistics. The context sections also provide relevant information on social, political, health, safety and other issues affecting women. = 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 analysis has a focus on gender equity and the extent the CP reached the most vulnerable populations, and findings that emerged from the different stakeholder groups are highlighted. However, the report could go further by having gender disaggregated output data (such as number of health service providers trained) and by including quotes that clearly bring out the voices of women and other rights holders. = 2</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Although the scope includes the examination on unintended effects, this is not explicitly addressed in the findings and is not reflected in the conclusions. One example, as noted above, was mentioned regarding the effect on household income when men are fined for domestic abuse but this was not framed as an unanticipated effect. = 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) Gender issues are woven into several recommendations, and there is a specific recommendation on interventions to address gender-based violence. = 3</p>
--	--

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totaling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory).

(**) Scoring uses a four point scale (0-3).

0 = Not at all integrated. Applies when none of the elements under a criterion are met.

1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

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	62	18	20	
Overall assessment level of evaluation report	Very Good			

Very good
very confident to
use

Good
confident to use

Fair
use with caution

Unsatisfactory
not confident to use

- (*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'Analysis and findings' has been assessed as 'Good', enter 40 into 'Good' column).
(b) Assessment level with highest 'total scoring points' determines 'Overall assessment level of evaluation report'. Write corresponding assessment level in cell (e.g. 'Fair').
(c) Use 'shading' function to give cells corresponding colour.

If the overall assessment is 'Fair', please explain

• How it can be used?

• What aspects to be cautious about?

Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory

Consideration of significant constraints

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:

Yes No

If yes, please explain: