

Organizational unit:	UNFPA Bangladesh Country Office	Year of report:	2020
Title of evaluation report:	Evaluation of the UNFPA 9th Country Programme of Assistance to the Government of Bangladesh		
Overall quality of report:	Good	Date of assessment:	2020 MAY
Overall comments:	This is a thorough evaluation that included an extensive data collection process with over 400 stakeholders participating, including beneficiaries. The methodology was well designed although data analysis processes could be more clearly described. The findings are detailed, predominantly address output-level results, and could be more analytical to clearly articulate the key results and accomplishments. The conclusions provide a good overview of the accomplishments and challenges of the programme and lead to a clear set of recommendations that should be useful for informing the next CP. The analysis of GEEW was well articulated but it could have been more clear how HRGE was mainstreamed throughout the evaluation process.		

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	<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
--	----------------------	-------------------	------

To ensure the report is comprehensive and user-friendly

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 mostly very well written, however there are a number of grammatical errors throughout (including incomplete sentences and the omission of determiners/articles such as 'the').
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)	Partial	The report, including Executive Summary, is over the maximum at 76 pages.
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	The normal structure is followed. There is not a separate section on Lessons Learned, nor is this required.
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?	Yes	The annexes are extensive. There are two volumes - one covers the required attachments (with the exception of information on the stakeholder consultation process but this is addressed in the report) and the other includes additional performance data, the ToCs, additional information on the UNFPA Coordination Role and Stakeholder Map.

Executive summary

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 mostly serves as a stand-alone section covering key points of the evaluation process and the main results. One issue is that it includes extensive use of acronyms, several of which are not explained (IP, FP, GE) and may not be understood by all of the target audience.
6. Is there a clear structure of the executive summary, (i.e. i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?	Partial	All components are included, and each has a bolded heading. However, the presentation would be more clear if subheadings for each criterion were used for Conclusions, and numbers were used for Recommendations. The Recommendations section is in paragraph form with a mix of complete and incomplete sentences making it somewhat difficult to follow.

7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	It is concisely written and is just over 4.5 pages in length.
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 main audience as well as the primary users of the evaluation are identified.
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	These issues are sufficiently covered in the Country Context and UNFPA Response & Programme Strategies chapters.
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 ToCs and results frameworks are provided for each CP component. The methodology section includes a brief assessment of the program theory (mainly noting that there was not an overall ToC but also highlighting other shortcomings) and indicates that there was not sufficient time to develop a comprehensive one within the evaluation timeframe.
<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 evaluation framework is referenced in the main text and is attached as Annex 4. It is comprehensive and includes all required elements.
5. Are the tools for data collection described and their choice justified?	Yes	Data collection processes are briefly outlined under the subsection called Data Sources and are initially listed as document review, individual and group interviews, KIs, informal group discussions, FGDs and direct observation. The rationale for selection appears to be based on the need to triangulate sources. Later in the subsection it is mentioned that there was also an attitudinal survey used with some stakeholders; there is not an explanation for how this was administered.
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)?	Yes	A stakeholder map is provided as an attached document. The consultation process is explained and appears thorough. It involved several rounds of consultation with CO staff and the ERG, including a validation workshop where findings and recommendations were presented and discussed.
7. Are the methods for analysis clearly described for all types of data?	Partial	Data analysis is only briefly addressed. Content analysis is noted as being used for interpreting qualitative data (but not further described), and consideration was given to looking at how marginalized groups were included in the CP design and implementation. The interpretation/analysis of quantitative data obtained through the attitudinal survey is not addressed.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	Limitations described include the lack of an overall TOC, ensuring a representative sample particularly in more remote areas, the interruption of the CP due to humanitarian crises arising, and the range of languages spoken by different stakeholders consulted. Mitigation strategies for each are presented.
9. Is the sampling strategy described?	Yes	The sampling strategy is outlined for the selection of sites for field visits and stakeholders consulted. A purposive approach was used for both and is adequately detailed. The limitations - primarily the level of representation of the samples - are noted and mitigation efforts are described.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	Evidence of this includes the disaggregation of evaluation participants both by stakeholder group and gender (presented in table form in the methodology section and in the annexed list of participants).
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	As noted above, the methodology section notes that the evaluators looked at how marginalized groups were included in the CP design and implementation, and these issues are also reflected in the evaluation matrix. The text includes several direct quotes from different stakeholders, most from beneficiaries.

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	Triangulation is noted as being applied during the analysis process, and there are several cases in the Findings where multiple sources are cited. The range stakeholders consulted as well as the extent of the sample size suggest that data collection was adequately robust.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Both types of sources are identified, and it is noted that most are qualitative. Within the findings, use of primary quantitative data is limited (the Likert Scale results from the attitudinal study is only referenced once, p 24) but secondary quantitative source data is presented more regularly (i.e. 13% decrease in pregnancy rates linked to the distribution of gift boxes, p 32) Reliability is discussed in respect to representativeness of stakeholders consulted and triangulation of data.	
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	Limitations and mitigation efforts are noted for both primary and secondary sources.	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	There is a subsection titled "Ethics and maintaining the quality of evaluation" which addresses issues of confidentiality and informed consent, and reference is made to following UNEG guidelines on integrating HRGE in evaluations, and to UNFPA's evaluation handbook. Beyond this, although there is not an explicit description of how representatives of beneficiaries participating in the evaluation were engaged or made to feel comfortable, further evidence of ethical considerations includes translators being engaged as needed who were native to those geographic regions.	
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 substantiated with evidence gathered in the evaluation process.	
2. Is the basis for interpretations carefully described?	Yes	The report includes substantial details and background information to show the basis for the findings.	
3. Is the analysis presented against the evaluation questions?	Yes	The questions are prominently stated at the beginning of each criterion section.	
4. Is the analysis transparent about the sources and quality of data?	Partial	Data sources are clearly cited in most sections. Reference is often made to the particular stakeholder group from which it was obtained. The evaluators also note where the findings are based on their own observations. However, sources are not consistently provided (for example, no sources are provided in the discussion on SRHR in Emergency Response in CXB, p 36-37).	
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The linkages are well presented in the results frameworks in Annex A-3 which details indicators, baseline, 2020 targets and achievements as of June 2019. General summary statements of achievements are given under Effectiveness (although sufficient to meet the criteria, the presentation would have been stronger if excerpts from the results framework - particularly comparisons of targets vs actuals - were more frequently given within the main report). Unintended effects are mentioned under Sustainability where two positive examples are highlighted including that child marriages appear to be decreasing in areas neighbouring targeted intervention communities.	
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	Attention is given to issues of accessibility to services for marginalized groups. There is little information on outcomes with the evaluators noting that there is limited data available within current administrative systems to identify inequalities across different marginalized groups, and that this was taken into account in the analysis process.	

7. Is the analysis presented against contextual factors?	Yes	Context is provided. This is particularly evident in the Effectiveness section where there are subsections on Facilitating Factors and Constraining Factors for each programme area.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The evaluation primarily looks at cross-cutting issues in terms of access of vulnerable populations and how CP services have targeted young women at risk, refugees and more remotely-located 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 clearly reflect the findings; the relevant EQs for each are provided in brackets.
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 effectively bring together the large amount of data presented in the Findings and enable a solid understanding of the CP accomplishments, challenges, underlying issues and potential directions for the next phase.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no indication of bias.
6. Recommendations		
	Yes No Partial	Assessment Level: Very good
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	The relevant conclusions for each recommendation are provided.
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations are mostly very clearly articulated. The intended users and suggested action plans are provided.
3. Do recommendations appear balanced and impartial?	Yes	They appear balanced by addressing both initiatives/approaches that should be continued and those that could be improved.
4. Is a timeframe for implementation proposed?	Yes	The timeframe is noted in the preamble to the section as being the next CP and in preparation for it.
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	They are given priority levels (of high and medium) and are presented in a way that facilitates a management response.
7. Gender		
	0 1 2 3 ^(*)	Assessment Level: Good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>		

<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>	<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) Although taken up in the evaluation, HRGE is not specified in the scope or objectives. = 1</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) HRGE is addressed - the evaluators note making a specific attempt to see if gender aspects were included in the design and implementation of the humanitarian interventions, and there is reference to vulnerable and marginalized groups under Relevance - but a more explicit focus on assessing GEEW in all component of the CP may have been useful. = 2 HRGE mainstreaming was done beyond humanitarian interventions. EQ1 refers to vulnerable and marginalized groups. The evaluation matrix in the annex included HRGE related assumptions under other EQs as well.</p> <p>2 c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) EQ8 considers the extent that UNFPA interventions and approaches address GBV and harmful practices in humanitarian settings. Otherwise, the investigation of GEEW can be seen in the indicators and interview protocols = 3 As mentioned above, assumptions reflected GEEW for EQs under relevance and effectiveness as well.</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 assessment of systems for collecting HRGE-related and disaggregated data is included within the evaluation framework under Relevance. = 3</p>
<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	<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) Evaluation participants are gender disaggregated however there is not a specific discussion on how the methodology is gender responsive. =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 are used and sample size of stakeholders met (400+) seems 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) A suitable range of sources is used and enabled triangulation, but description of analysis of convergence of sources of data not discussed. Validation processes were carried out. = 2</p> <p>2 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 range of stakeholders consulted included beneficiaries but it is not clear if they were representatives of vulnerable groups. = 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) The methodology included a subsection on ethical practices and highlighted that UNEG and UNFPA guidance was used, and that confidentiality and informed consent were obtained. However, it could have gone further by specifying practices such as timing/locations of interviews, protocols when abuse is disclosed, etc. = 2</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 chapter discusses the relevant instruments/policies in place and the HRGE-related challenges. = 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 show the perspectives of different stakeholder groups and include several direct beneficiary quotes = 3</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) 3 As noted above, two examples of unanticipated HRGE-related effects are described. = 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 that address GEEW issues including the need for more gender disaggregated data, for additional focus on marginalized groups, for improving effectiveness of GBV initiatives, etc.= 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	33	47	20	
Overall assessment level of evaluation report		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: