Organizational unit:	**UNFPA**	**Year of report:**	**2017**
Title of evaluation report:	**UNFPA Country Programme Evaluation: Myanmar CP3 2012- 2017**		
Overall quality of report:	**Very Good**	**Date of assessment:**	**7 September 2017**

Overall comments: The Myanmar Indepent Evaluation report is strong: It is mostly well-written in English (although quality is varied in sections) and clear. The evaluation framework is clearly described in the Methodology section and supporting annexes. The analysis is transparent about sources and quality of data - and the evaluators were diligent in ensuring data were consistently cited, reliable and valid. The cause-effect links between outputs and outcomes are rationally explained. Conclusions follow logically from the analysis and present important issues that UNFPA needs to consider in decision-making and next cycle program development. Recommendations are action-oriented, however they could be better prioirtized and, in some cases, more clearly written. GEEW is integrated in the evaluation scope of analysis and the evaluators collected and presented GEEW-related data.

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:	**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	1. The report was mostly easy to understand although the quality of writing varied in different sections. Some mixing of past and current tense, and some incomplete sentences could have been addressed by light editing.	
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	2. The report is is over the reasonable length for a Country Programme Evaluation at 101 pages, 21 pages over the normal limit.	
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	3. The report is logically structured.	
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	4. The annexes include: the ToR; a list of interviewees and documents consulted; the evaluation matrix and individual programme theories of change for SRHR, P&D, Gender and Humanitarian Setting; and an additional discussion of findings. While methodological tools are not specifically provided, Annex 6 does provide additional rationale and details on the data collection methods used and the stakeholder consultation process.	
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?	Yes	5. The Executive Summary serves as a stand-alone section and presents the main findings of the evaluation.	
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)?	Yes	6. The Executive Summary is organized in a clear manner, presenting the purpose, objectives, methodological approach, conclusions and recommendations as required.	
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	No	7. At 7 pages, the executive summary exceeds the maximum length. It could have been more concisely formulated if the sections on Country Context and Limitations of study were left to the full report.	

2. Design and Methodology	Yes No Partial	Assessment Level:	**Fair**
To ensure that the evaluation is put within its context			
1. Does the evaluation describe the target audience for the evaluation?	Yes	1. The target audience of the evaluation are listed as the UNFPA CO, APRO, Headquarters, Government partners and other UN agencies and development partners.	

2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	2. The country context is extensively described. This section is well written but, at 11 pages, appears rather excessive for an evaluation report. The institutional context is also thorough. It describes the outcomes and outputs of each program area.
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	3. The report discusses limitations in each programme's theory of change. In addition, the limitations were discussed extensively with programme staff and addressed through the results matrix and evaluation methodology. Mainly, the evaluation team found that there were too many activities with unclear links to outcomes. In addition, they found some of the objectives and outcomes to be too similar. While there is adequate discussion of the intervention logic, and it seems to have been addressed with programme staff, there is no description of the reconstructed theory of change. Instead, the need to adjust program priorities and "accompany with theories of change" is provided as a recommendation (Recommendation 1).
To ensure a rigorous design and methodology		
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	4. The evaluation framework is provided in the annex and briefly mentioned in the main report. The framework carefully details the assumptions, indicators and data sources and collection methods for each evaluation question.
5. Are the tools for data collection described and their choice justified?	Yes	5. Some details on the chosen data collection tools are provided in the text of the report, however more substantial information on the tools and justification for their use are provided in Annex 6.
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	6. The stakeholders are identified and the consultation process was described as including regular checks with program staff, and a final meeting where all parts of the report, including the recommendations, were shared.
7. Are the methods for analysis clearly described for all types of data?	Partial	7. The methodology section mentions that "the methods for data collection and analysis are determined by the type of evaluation questions formulated" (p.11) but does not specify the analysis processes used. The more detailed discussion on methodology in Annex 6 notes that there was extensive consultation with stakeholders about data collection and interpretation but also does not identify any type of analytical method.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	8. Limitations of the methodology are clearly explained and referenced in terms of how they affect the findings.
9. Is the sampling strategy described?	Yes	9. Purposive sampling was applied and the limitations of this approach were discussed and minimized in consultations with stakeholders. The evaluation team utilized purposive sampling based on programme interventions, beneficiary populations and the characteristic of geographic locations in order to deliberately ensure representation. The criteria for selecting field visit sites are clearly explained and used to help minimize selection bias (p.22).
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	10. Disaggregaged data was provided for the overall number of evaluation participants.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	11. The Methodology is appropriate for assessing cross-cutting issues, they specifically did well to address Gender in the methodology.

3. Reliability of Data	Yes No Partial	Assessment Level:	**Very good**
To ensure quality of data and robust data collection processes			
1. Did the evaluation triangulate data collected as appropriate?	Yes	1. The evaluation made clear attempts to triangulate data methods, including desk reviews, observations, site visits, focus group discussions, and interviews. In addition, data was validated through exchanges with the Country Office programme officers and ERG members in order to support data validity.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	2. Data sources were also clearly identified and diverse, including stakeholders from the UNFPA Country Office and other UN agencies, line ministries, international and national NGOs, academic institutions, faith groups, donors and other beneficiaries. The evaluation team consulted 254 stakeholders, out of which 65% were female. Detailed information about qualitative and quantitative data sources is provided in the Annexes and referenced throughout the report. At the same time, the evaluation report cited significant shortcomings in the data available, making it appear like they didn't have sufficient data despite the diverse data collection methodologies used and stakeholders consulted: "The short timeframe of this evaluation did not allow the team to collect primary quantitative data for related areas and accessing data with some consistency was a major challenge. Reliability as well as the lack of secondary data was a limitation in general. In order to mitigate these limitations, the evaluation team collected qualitative data where secondary data was limited or absent. In consultation with Country Office, the evaluation team sought for numerous data sources to report what is appeared to be most reliable. " (p 14)	
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	3. Information about limitations and mitigation measures is presented on p. 20.	

4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	4. Ethical considerations during data collection are provided in the annex but would have been better placed in the main report.

4. Analysis and Findings	Yes No Partial	Assessment Level:	**Very good**
To ensure sound analysis and credible findings			

1. Are the findings substantiated by evidence?	Yes	1. There is substantial evidence presented for the findings, including a supplementary detailing of findings in the annex.
2. Is the basis for interpretations carefully described?	Yes	2. The basis for interpretations is carefully described in the findings, rooting analysis in data collected - with specific cases being referenced to highlight interpretation. While this is done well throughout the report by referencing specific quotes from interviews and focus group discussions or citing specific douments, there are several instances where more specific data could be referenced. For example, the following finding could have been strengthened with mention of the specific amount of resources mobilized (if available): "UNFPA played a leading role in mobilizing the external resources for the census. UNFPA has also contributed by providing a permanent resident Chief Technical Advisor and mobilizing a significant amount of resources from donors group."
3. Is the analysis presented against the evaluation questions?	Yes	3. The evaluation questions are listed prior to the relevant findings, as are a concise summary of findings.
4. Is the analysis transparent about the sources and quality of data?	Yes	4. The analysis transparently presents the sources of data, and discussion of the quality of data is included througout and in the 'limitations' section.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	5. The report discusses the cause and effect links between the intervention and its end results in the body of the text. However, drawing cause and effect links was also cited as a key limitation due to several changes in indicators and programme activities over the course of the programme period (2012 - 2016)(p. 59). This was a finding in itself, and recommendations were made to improve the programme logic in the recommendations section. The evaluation does well to mitigate the effects of weak cause-and-effect links in the programme theory with qualitative data from beneficiaries supporting intended and unintended outcomes.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	6. The analysis presents/disaggregates outputs and outcomes for different stakeholders.
7. Is the analysis presented against contextual factors?	Yes	7. The evaluation reports provides description of contextual factors in detail, and situates findings and analysis in those contexts.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	8. There is a very thorough and strong analysis of equity, vulnerability and gender issues

5. Conclusions	Yes No Partial	Assessment Level:	**Very good**
To assess the validity of conclusions			

1. Do the conclusions flow clearly from the findings?	Yes	1. The conclusions emerge from the findings.
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	2. They are well formulated, effectively summarizing and presenting the larger context for the main findings.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	3. Evidence of non-bias includes that the evaluation team incorporated a debriefing session on the preliminary results for Country Office staff in order to validate the findings and test tentative conclusions. Feedback was incoporated into the conclusions.

6. Recommendations	Yes No Partial	Assessment Level:	**Fair**
To ensure the usefulness and clarity of recommendations			

1. Do recommendations flow logically from conclusions?	Yes	1. The report indicates which conclusions inform each recommendation.

2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	2. They are mostly well formulated and targeted, although there are some editing issues that affect readability. Some recommendations could be worded more specifically. For example, #2 suggests establishing "a national capacity development plan that is based on identified gaps to enhance coordination and harmonization and to avoid overlaps" - it is not clear, even from the accompanying text, what the basis of "overlaps" is and what capacity building initiatives this refers to. The use of 'enhance' contributes to the vagueness. Furthermore, financial and technical implications are not addressed.
3. Do recommendations appear balanced and impartial?	Yes	3. They appear to be balanced and impartial.
4. Is a timeframe for implementation proposed?	Partial	4. The preamble to the recommendations suggests that the recommendations should be considered for the development of the SP 2018-2021. However, there are no timeframes proposed for individual recommendations.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Partial	5. All recommendations were classified as 'high priority'. The sub-recommendations are sufficiently well formulated for management response and follow up.

7. Gender	0 1 2 3	Assessment Level:	**Good**
To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)			
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	1. The scope of the evaluation integrates GEEW as GEEW is a programme component itself and cross-cutting issue. Indicators have gendered components due to their focus on sexual and reproductive health and gender-based violence but do not incorporate GEEW dimensions specifically (for example, no indicators require the evaluators to disaggregate data by gender). However, data is consistently disaggregated by gender, when applicable.	
2. Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved?	3	2. Gender was built into the evaluation questions, especially dealing with relevance and effectiveness. Their methods incorporated gender factors and there was an evident application of gender in the evaluation process. The conclusions and recommendations incorporate gender factors.	
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	3. As described above, the methdological approach was gender-responsive. However, the methodological tools are not provided so it is not clear whether those, specifically, are gender-responsive.	
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	4. All three sections include substantial consideration of gender issues, including GBV. Gender-based violence and the absence of national gender disaggregated census data are clear problems in the country, and the evaluation team has acknowledged this through their findings and associated recommendations.	

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

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	7	31	
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

While the evaluation was too long, it covered most of the areas successfully.

Consideration of significant constraints

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

If yes, please explain: