

Organizational unit:		Year of report:		2018									
Title of evaluation report:		EVALUATION OF THE UNFPA 7TH COUNTRY PROGRAMME OF ASSISTANCE TO THE PHILIPPINES - Final Report											
Overall quality of report:		Very Good		Date of assessment:									
				31 July 2019									
<p>Overall comments: This evaluation was well designed and clearly presented. It provided a good overview of the Country Programme's accomplishments and strengths for the three island groups covered (parts of Luzon, Visayas and Mindanao), and included a comprehensive set of recommendations for consideration for the next planning cycle. The evaluation clearly captured the outputs of UNFPA's contribution and, in part, how these contributed to achieving (some) outcomes by use of a carefully designed theory of change for each of the three specific programmes. Data was collected through document analysis, collection of primary and secondary data from the country health system, structured and unstructured interviews, field visits as well as direct observation, and focus group discussions using a purposive sample. Triangulation of sources, methods tools and data were made to validate findings. The inclusion of factors that facilitated and hindered the country programme, the delineation of unintended effects (both positive and negative), the identification of good practices as well as lessons learned are directed toward informing the design of the next country programme (which will focus on another island and have greatly reduced funding resources). The conclusions flow clearly from the findings and lead to a set of prioritized recommendations directed to the design of the next country programme (already underway).</p>													
<p>Assessment Levels</p> <table border="0"> <tr> <td>Very Good</td> <td>strong, above average, best practice</td> <td>Good</td> <td>satisfactory, respectable</td> <td>Fair</td> <td>with some weaknesses, still acceptable</td> <td>Unsatisfactory</td> <td>weak, does not meet minimal quality standards</td> </tr> </table>						Very Good	strong, above average, best practice	Good	satisfactory, respectable	Fair	with some weaknesses, still acceptable	Unsatisfactory	weak, does not meet minimal quality standards
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 <u>assessment level</u> followed by main <u>comments</u> . (use 'shading' function to give cells corresponding colour)										
1. Structure and Clarity of Reporting			Yes No Partial	Assessment Level:	Very good								
<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?			Yes	The report is clearly written. There are a few grammatical, spacing and punctuation errors (i.e., first two paragraphs on p. 36) but this does not affect the readability of the report.									
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 main report is 70 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 report structure is logical, following the recommended structure, and clearly distinguishes between findings, conclusions, and recommendations. Lessons learned and factors facilitating and hindering programme performance are interspersed in response to the third specific evaluation objective which was to "draw key lessons from past and current cooperation" (p. 6).									
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 and go beyond the minimum to include the theory of change, logical framework, and evaluation data. The table of contents lists some, but not all of the annexes (the last 5 are grouped under the heading "Other Supporting Documents"; it would be useful if these were all listed in the Table of Contents).									
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	The executive summary is a well-written overview of the evaluation and its main results. It serves as a stand-alone section.									
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	The structure is clear and includes all sections.									
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?			Yes	It is 5 pages long.									
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 audience is specified as part of the purpose and objectives of the evaluation.									
2. Is the development and institutional context of the evaluation clearly described and constraints explained?			Yes	There is a chapter on Country Context. It clearly outlines development challenges in respect to SRHR, population, and gender equality, as well as national strategies to promote gender equality. There is also a chapter on UNFPA's response and programme strategies. It addresses the institutional context, specifically the evolution of the country programme, its geographic coverage and how intervention areas were selected, as well as partnering agencies.									
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 evaluation reconstructed the intervention logic and theory of change. The shortcomings of the existing theories of change were explained - including (a) a lack of clarity on the links between outputs and outcomes and (b) missing assumptions (i.e. the assumptions were not included).									

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	The evaluation framework is incorporated into the evaluation matrix (presented in the annex).The framework can also be inferred from the text, including the evaluation questions (as part of Evaluation Criteria and Evaluation Questions).The evaluation matrix includes all information listed here, including the evaluation questions, assumptions, indicators, data sources and methods for data collection.	
5.Are the tools for data collection described and their choice justified?	Yes	The types of tools used are listed and the interview protocols for different stakeholder groups are provided in the appendix.The rationale for selecting the particular methods/tools was explained, and the evaluation noted that data from various sources would be triangulated.	
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	While reference is made to a "stakeholders map" in the text (p 7) and in Annex 1 (p 6) that the evaluation team prepared and used to identify sources of information, it is not included in the report. The overall stakeholder consultation process is not described per se, although the team notes that "the evaluation will adopt an inclusive approach" (Annex 1, p 6) and specific mention is made of the Evaluation Review Group (p 11), as well as particular aspects of the consultative process (p 110). Annex 1 does delineate the roles of specified stakeholders (pp 11-12) in the evaluation which points to their engagement in the consultation process.The report indicates which stakeholders will be involved in the review of "the draft report" for quality assurance (p 11 and Annex 1).	
7.Are the methods for analysis clearly described for all types of data?	Partial	The report provides only a general description of the approach to data analysis. It notes that content analysis was used but there is no further information on how this was done, or any mention of other types of analysis.	
8.Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The limitations and risks to the evaluation are clearly articulated, i.e. the size of the country, the spread of programme intervention, and limited data availability. Actions to mitigate/reduce these are also provided.	
9. Is the sampling strategy described?	Yes	Purposive sampling was employed for site selection, and the criteria for selecting each is given in table form. A 'convenient' sample (which is presumably 'convenience' sampling) was used to select beneficiaries for focus group discussions. A purposive sampling approach was used for interviews; the limitations of this approach and mitigation actions are noted.	
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	A table detailing the range of stakeholder groups consulted is provided and the respondents are disaggregated by gender.	
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	While the evaluation report does not include a specific section on whether or how the evaluation methodology itself is appropriate for assessing the cross-cutting issues (or gender and human rights responsive), an evaluation question explores whether the country programme integrated a gender responsive and HR-based approach to planning and implementation. It is also noted that the evaluation employed a participatory approach in that respondents were given the opportunity to freely discuss the programme and propose program improvements to better meet community needs. Taken together, these suggest that the methodology is, to an extent, appropriate for assessing cross-cutting issues.	
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	The report specifically states that "a systematic triangulation of data sources, ... methods... and tools" was done and that the results of the evaluation will be validated by two workshops with the evaluation reference group and CO staff as well as by a broader workshop held with other relevant stakeholders to present the findings and recommendations. Data is triangulated, document sources are regularly footnoted, and qualitative findings are regularly attributed to specific stakeholder groups.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Yes, the evaluation used both types of data. Quantitative data primarily comes from secondary sources. The reliability of the data is addressed. The evaluators note limitations with regard to data, in both the methodology of the report as well as in the findings (i.e. issues with the availability of national, local and regional data are taken up under Observations and Challenges on p. 42).	

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	There is a section on data gaps that addresses gaps in secondary data (national data collection processes that would enable the assessment of changes in the development conditions of specific target populations) and how that limitation was addressed. Limitations and mitigation actions were also briefly described for primary data sources (interviews and sites visited). The report states that the "the evaluation team will use a variety of validation mechanism" and, as noted above, specifically states that "a systematic triangulation of data sources ... methods... and tools" are used and that the evaluation will also be validated by two workshops with the ERG and CO staff, as well as by a broader workshop held with others.
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 on "Ethics and maintaining the quality of evaluation" where precautions taken to protect respondents' rights are explained, i.e. obtaining informed consent and ensuring stored data does not have identifiers.
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	The evaluators have been careful in substantiating findings with evidence throughout.
2. Is the basis for interpretations carefully described?	Yes	This is also well done throughout, and is particularly evident in the "Effectiveness" section where baseline and end-of-project data is compared.
3. Is the analysis presented against the evaluation questions?	Yes	The relevant evaluation questions are listed at the beginning of the discussion of each criteria, and the analysis is structured accordingly.
4. Is the analysis transparent about the sources and quality of data?	Partial	Data sources for findings (i.e. source documents, stakeholder group) are usually cited, but the evaluators could have been more methodical in doing this (i.e. on p. 44, there is reference to the number of teachers benefiting from GRM training but the source of this data is not given). In a number of cases dates are not provided for the information presented (i.e. Figure 7 on p. 35 is cited as coming from the "Zuellig Family Foundation" but not when). In some cases, the source could be more specific (i.e. for Figure 6 on p. 34, the source is only listed as "Provincial data").
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The framework for assessing causal connections was well established in the evaluation matrix, and the links between the programme activities and the results are evident throughout the text of the findings. Figure 7 provides a strong visual example of the connections between the leadership/governance activities and improved local health outcomes. The outputs of each component of the country programme are clearly delineated in tabular form and elaborated in the narrative text. While a direct cause-effect linkage between outputs and outcomes was not established as such, outcomes were often identified in terms of specific improvements that were achieved, or in a sequence of interventions that resulted from these outputs. Contribution was also established by way of quotes about the benefits that resulted, or by citing models of best practice that could be emulated by others, and by noting the progression in the evolution of the programme over time. Unintended effects - positive as well as negative - were identified.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The country programme has projects that target specific groups (indigenous groups, other minorities, youth) and the evaluation examines outcomes for these different groups.
7. Is the analysis presented against contextual factors?	Yes	The analysis was presented in terms of contextual factors including the evolution of the programme over time, by "good practices", by lessons learned that have broader implications for the programme as a whole, and by a discussion of the facilitating and hindering factors that influenced implementation of CP7. Contextual factors are frequently cited (i.e. the effect of natural disasters on program activities and UNFPA's response, the very high prevalence of HIV/AIDS in the country, the strategic shift in UNFPA programming). Another example is found under effectiveness where contextual factors are provided as part of the analysis on the extent of indicator achievement.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	Comments on the extent to which vulnerable groups have their voices heard are found under Relevance. Cross-cutting issues are also addressed under Effectiveness where there is a sub-section on Gender Equality which, among other issues, discusses UNFPA's work in promoting the participation of women and girls who have disabilities in decision-making in RH and GBV support systems.
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 are clearly drawn from the findings. The evaluators have specified the respective evaluation question numbers that are linked to 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?	Yes	The conclusions go beyond the findings to provide a comprehensive picture of the accomplishments of the CP and issues to build upon.

3. Do the conclusions appear to convey the evaluators' unbiased judgment?	Yes	There is no indication of bias. The dissemination workshops with key stakeholders was a means of checking for and receiving feedback on any potentially biased judgment by the evaluators.	
6. Recommendations	Yes No Partial	Assessment Level:	Very good
To ensure the usefulness and clarity of recommendations			
1. Do recommendations flow logically from conclusions?	Yes	Recommendations, organized at the strategic and programmatic (operational/management) levels, flow from and are built upon the conclusions.	
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 very clearly stated, and are action-oriented. Although the financial implications are not provided for each recommendation, the preamble to this section states that the evaluators took into account the significant reduction in the core budget for the subsequent CP.	
3. Do recommendations appear balanced and impartial?	Yes	These are well-rounded and there is no indication of bias. The presentation of evaluation results for the in-country dissemination workshop was intended to validate findings and ensure recommendations were appropriately balanced and impartial.	
4. Is a timeframe for implementation proposed?	Yes	Yes, recommendations were meant to feed into and shape the next country programme (CP8), currently under development.	
5. Are the recommendations prioritized and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	The recommendations are adequately prioritized and are clearly presented at both the strategic and programmatic levels.	
7. Gender	0 1 2 3 (**)	Assessment Level:	Very 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?		While the evaluation does not include a specific objective dedicated to gender equality, gender equality is mainstreamed across evaluation criteria, including how relevance, effectiveness, efficiency and sustainability were conceptualized and applied. Moreover, there is a specific evaluation question on the extent to which the country programme integrated a gender responsive and HR-based approach to planning and implementation. The evaluation also assesses the extent to which the CP collected sufficient information during its implementation to be able to measure progress on human rights and gender equality, and limitations/shortcomings are noted.	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		The methodology is not explicit about how gender was considered. However, the evaluation employed a mixed methods approach, collecting both quantitative and qualitative data, which allowed for the assessment of the extent to which gender equality was advanced by the country programme. Gender-disaggregated data is collected and used, a diverse range of stakeholders participated in the evaluation, including beneficiaries, and data is triangulated. Finally, the evaluation report has a sub-section that addresses how ethical standards were upheld.	
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		The evaluation has a background sub-section on the relevant national strategies and the policy environment vis a vis gender equality. Human rights issues, such as the rights of migrant workers and indigenous rights, are also discussed in the sub-section. While the findings include data analysis that explicitly triangulates the voices of different groups, this is not consistently done. Disaggregated quantitative data is used, but there is, for example, minimal use of quotes or stories to highlight the voices of different groups. Unanticipated effects of UNFPA support on human rights and gender quality are discussed under the evaluation criteria sustainability, where, for example, it is noted that a range of vulnerable groups that were not initially targeted by UNFPA benefited unexpectedly from UNFPA's humanitarian work. Finally, there are several GEEW-related recommendations including ones that address gender based violence, and the need for more gender-disaggregated data to be collected.	
(*) 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 totalling 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			
		Assessment Levels (*)	

Quality assessment criteria (scoring points*)	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	47	40	13	
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
<p>(*) (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.</p> <p>If the overall assessment is 'Fair', please explain</p> <p>• How it can be used?</p> <p>• What aspects to be cautious about?</p> <p>Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory</p> <p>Consideration of significant constraints</p> <p>The quality of this evaluation report has been hampered by exceptionally difficult circumstances: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>				