

**Organizational unit:** \_\_\_\_\_ **Year of report:** **2018**

**Title of evaluation report:** **EVALUACIÓN FINAL DEL VII CICLO DE PROGRAMA PAÍS GUATEMALA 2015-2019**

**Overall quality of report:** **Very Good** **Date of assessment:** **12 April 2019**

**Overall comments:** This evaluation provides a thorough analysis of the country programme. While improvements could be made in the style of the evaluation report as well as in detailing the context in which the evaluation is undertaken, the findings were carefully developed by the evaluators, who effectively used the data to ground the results observed. Carefully developed conclusions and recommendations emerged from the findings, to be used in the development of the next country programme by UNFPA and its related stakeholders.

**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 <u>assessment level</u> followed by main <u>comments</u> . (use 'shading' function to give cells corresponding colour)		
<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure the report is comprehensive and user-friendly</i>			
<b>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?</b>	Yes	Though the report is quite long, it is easy to read.	
<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>	Partial	The CPE is 75 pages not including glossary and index.	
<b>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)?</b>	Yes	The report is structured in a logical way - and follows the expected structure.	
<b>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?</b>	Yes	The annexes contain the required information.	
<i>Executive summary</i>			
<b>5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?</b>	Yes	The summary is presented as a standalone section and includes the main results of the evaluation.	
<b>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)?</b>	Partial	While the summary includes all the main subsections, the manner in which the conclusions and recommendations are presented (in a two-column, very dense/condensed table) impacts clarity/readability and complicates the ability to connect/relate the two areas.	
<b>7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	No	The summary is eight pages.	

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure that the evaluation is put within its context</i>			
<b>1. Does the evaluation describe the target audience for the evaluation?</b>	No	While the target audience is described in the terms of reference to the evaluation (which is annexed to the report), there is no mention of the target audience in the design/methodology of the report.	
<b>2. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	There is a comprehensive description of the context, including the institutional context of UNFPA as an organization.	
<b>3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?</b>	Partial	There is no description of the intervention logic in the text of the report itself. An evaluation matrix was developed, however, suggesting that the evaluators were implicitly aware of the theory of change underlying the programme.	
<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 described in the text and the evaluation questions are described in both the text and, in considerable detail, in Appendix 4. The matrix includes detailed information on indicators, data sources and collection methods.
5. Are the tools for data collection described and their choice justified?	Yes	Yes, data collection methods were described (i.e. the evaluation primarily used interviews and focus groups discussion, as well as documentary review for quantitative and background data) and the rationale for their selection was provided (they were chosen based on the questions that were to be answered particularly vis a vis beneficiary groups).
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 a comprehensive stakeholder mapping is not described as such, on p. 21 it is stated that the team created one. A Reference Group was constituted and consulted throughout the process, including on the draft recommendations.
7. Are the methods for analysis clearly described for all types of data?	Yes	The analysis methods were described in the section on Metodología y Proceso. For example, the section on relevance noted how various documents were analyzed. Additionally, the following was noted, illustrating an approach to analysis: "A partir de una primera revisión documental sobre los principales avances del programa se identificaron y analizaron otros elementos y/o factores que explican cómo y por qué las acciones de UNFPA contribuyeron, o no, a los efectos esperados. Para ello, se reconstruyeron hipótesis de trabajo que fundamentaron las acciones en los distintos componentes del Programa de País. Esta reconstrucción se realizó con la información obtenida en las entrevistas al personal UNFPA, socios, actores y destinatarios del programa. Los hallazgos y conclusiones generales permitieron validar las hipótesis."
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	No	There was little discussion of limitations.
9. Is the sampling strategy described?	Yes	A purposive sample was used and the UNFPA country office helped determine who were "direct or indirect partners" (socios directos e indirectos).
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	There was considerable disaggregation, particularly of data acquired through document reviews.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The evaluators note the importance of gender and human rights as well as vulnerability (including in the political context and the consequences of poverty) and include cross-cutting issues as part of the evaluation design: "El equipo de evaluación ha adoptado el Enfoque de gestión por resultados, de derechos humanos y de género, para valorar la contribución del UNFPA al logro de los outputs y outcomes del CPAP. Tomando en cuenta que, por la tipología del programa, sus temas, instituciones del estado y organizaciones de la sociedad civil involucradas, el rol de UNFPA no es de brindar respuestas directas a los grupos metas mediante servicios y productos, sino de apoyar las instituciones del estado en cumplir sus obligaciones como garante de los derechos y a la sociedad civil como actores clave para la promoción de las instancias de los titulares de los derechos, que representan los grupos metas finales del CPAP: mujeres, niñas, adolescentes, especialmente de los grupos en mayor vulnerabilidad, como los de los pueblos indígenas, personas con VIH-SIDA, personas con discapacidad." In data collection there was an effort to ensure that the groups (especially women) were included in interviews in the field and in the on-line survey.

<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure quality of data and robust data collection processes</i>			
1. Did the evaluation triangulate data collected as appropriate?	Yes	There was clear evidence of triangulation, where multiple sources of data were used (i.e. document review, interviews and the survey) together to reach findings.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation used quantitative data, especially financial but also from the survey, for findings as well as an extensive use of qualitative data from interviews.	
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?	Partial	There was little discussion of limitations in the very detailed finding analysis.	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	While the report does not explicitly discuss how the evaluators ensured that data was collected ethically and non-discriminatory, data was acquired and presented anonymously, suggesting compliance with minimal ethical standards.	

<b>4. Analysis and Findings</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	Each finding is based on extensive (and extensively described) data.	

2. Is the basis for interpretations carefully described?	Yes	In each case, the basis for the interpretation is clear, especially in the findings related to effectiveness.
3. Is the analysis presented against the evaluation questions?	Yes	All of the analysis is organized around the evaluation questions that are themselves organized around criteria.
4. Is the analysis transparent about the sources and quality of data?	Yes	Yes, the evaluators are transparent about the sources of data used, as well as the quality of the data.
<b>5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</b>	Yes	The links are usually shown, although what are called output (productos) are often what is generally considered outcomes. There were some unintended outcomes and these were noted.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis is careful to note differences by target group. One that was given attention were LGBTI persons.
<b>7. Is the analysis presented against contextual factors?</b>	Yes	The analysis is very clear about the contextual factors. For example, in one case the fact that an election led to a major turnover in counterpart personnel was a factor affecting the result.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	There was an extensive analysis of gender equality in most cases, however human rights were less elaborated upon, but were considered in the analysis of UNFPA's strategic role in Guatemala, where UNFPA is considered the guiding institution on reproductive health and rights.

<b>5. Conclusions</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
-----------------------	----------------------	-------------------	------------------

*To assess the validity of conclusions*

<b>1. Do the conclusions flow clearly from the findings?</b>	Yes	Yes, they clearly flow from and summarize the findings. Both the findings and the conclusions are organized by evaluation criteria (relevance, effectiveness, efficiency, sustainability, etc.).
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 summarize the findings in a way that illustrates an understanding of the broader context and the findings interplay with the context and institutional factors.
3. Do the conclusions appear to convey the evaluators' unbiased judgment?	Yes	There is no evidence of bias.

<b>6. Recommendations</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
---------------------------	----------------------	-------------------	-------------

*To ensure the usefulness and clarity of recommendations*

<b>1. Do recommendations flow logically from conclusions?</b>	Yes	The recommendations are structured by criteria and respond to conclusions and the findings upon which they are grounded.
<b>2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b>	Partial	The recommendations are clearly written and most seem to be directed to the UNFPA country office. In some cases, a full description of what should be done, how and with what resources is included, but in others, only what should be done is provided, without details on the human/financial and/or technical implications.
3. Do recommendations appear balanced and impartial?	Yes	There is no evidence of bias.
4. Is a timeframe for implementation proposed?	Yes	The recommendations are for the next country programme which should be under development now.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Partial	While the recommendations are clearly presented, they are not prioritized.

<b>7. Gender</b>	0 1 2 3	Assessment Level:	<b>Very good</b>
------------------	------------------	-------------------	------------------

*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?	3	The scope of analysis and indicators of the evaluation integrate gender equality and women's empowerment. For example, as articulated in the objectives of the evaluation, the evaluation will assess the extent to which gender and human rights are integrated in the country programme.
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	Most of the questions in the evaluation are general and the GEEW dimension is more implicit than explicit, except where the programme element is gender specific. Three of the eleven questions refer specifically to gender. In this sense, the evaluation criteria and questions addressed GEEW in a satisfactory way.

3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	There is an effort to include gender-responsive tools in the questionnaires and in the document analysis. Gender is included as a variable in analyzing the contribution of UNFPA support in some areas, such as working with young boys. The evaluation was designed in a way to capture, for example, the country programme's work on new concepts and approaches that are relatively innovative such as an examination of "new masculinities".
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	Gender is incorporated into the findings, conclusions and recommendations where relevant. It is found in two of the six conclusions including one that is specifically on gender-based violence.

(\*) 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			
<b>Total scoring points</b>	<b>58</b>		<b>42</b>	
<b>Overall assessment level of evaluation report</b>	<b>Very Good</b>			
	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 there were elements of the evaluation that had some problems, on the whole, through the findings, conclusions, recommendations and lessons learned, the evaluation as a whole was rated as very good (with 51 points).

#### Consideration of significant constraints

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

Yes  No

If yes, please explain: