

Organizational unit: _____ **Year of report:** _____

Title of evaluation report: **Evaluación del Programa de país de Nicaragua 2013-2017**

Overall quality of report: **Fair** **Date of assessment:** **10 March 2018**

Overall comments: The evaluation covers a period of time in which changes in government priorities, funding and human resources affected the UNFPA country programme. The evaluation methodologies used were standard, but how they were applied is often not well-described. In some cases, why areas/sites were visited is not clear. The conclusions suggest a need for improvement in the next cycle and specific recommendations to that effect are provided. There was a strong effort to include gender analysis in the evaluation.

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
<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	It is generally easy to understand, although there are a few problems with grammar and the way in which acronyms are presented, especially in the executive summary. Additionally, acronyms are not always clearly introduced.
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	The main body of the report exceeds the prescribed reasonable length at 83 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	Yes, it is structured in a logical way, following the standard structure.
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	While the annexes contain some of the information, the annexes do not include, for example, the methodological tools used in the evaluation.
<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?	Yes	The executive summary contains all of the required information to be considered 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)?	Partial	The description of the intervention was not well-presented and the methodology was not well-described (it was only one sentence).
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The executive summary was six pages.

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 evaluation is for accountability and learning, informing UNFPA and its partners in planning for the next programme cycle of 2019-2023 (p16).	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Partial	The general/broad context in which the evaluation is undertaken is described. However, the development and institutional context is only briefly described in its framing of Nicaragua's country programme activities in the "orange" quadrant - focusing on development activities such as capacity development and knowledge management.	
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	The evaluators note that they needed to produce a revised theory of change that would indicate outcomes. However, how it was reconstructed is not clear and the evaluation matrix in the Annex only shows the hypotheses being tested. In the text, the evaluators use the term output (producto) to mean outcome (the change that UNFPA actions are expected to/intended to contribute to). What UNFPA does, specifically, is not clear in the description.	
<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 questions are clear (and were further refined from the ToR) and indicators are shown, as well as where and how data were collected.	
5. Are the tools for data collection described and their choice justified?	Yes	The main tools were documents, interviews, focus groups and observation and these were both described.	
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 evaluators consulted 97 stakeholders; a description of stakeholders consulted is provided and disaggregated by gender (p22), but there is not a comprehensive stakeholder map. In addition, the evaluation team briefly describes stakeholder consultation (p18).	
7. Are the methods for analysis clearly described for all types of data?	Yes	Analysis is discussed, though only briefly.	
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Partial	There was a section on evaluability that noted limitations and mitigation strategies, in general terms. There is some discussion of limitations, mostly surrounding inconsistencies in indicators, but no significant discussion of how potential bias in sources will be overcome.	
9. Is the sampling strategy described?	Partial	Who was to be sampled is described as an illustrative sample, but how specific informants were chosen (i.e. the approach) is not described. Two field visits were made (Bluefields and Matagalpa) but why these were chosen is not described in the methodological section.	
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	Disaggregated data were obtained, particularly relating to gender.	

11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Partial	The evaluators conducted a full review of how gender and human rights would be incorporated into the evaluation planning and analysis strategy; however how gender and human rights would be included in sampling or data collection (e.g. would separate focus groups with male/female or minority groups be conducted?) was not discussed.
--	---------	--

3. Reliability of Data	Yes No Partial	Assessment Level:	Fair
-------------------------------	----------------------	-------------------	-------------

To ensure quality of data and robust data collection processes

1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation sought to triangulate data across primary and secondary sources, including quantitative data from reports with qualitative data from interviews.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	The evaluation did not consistently identify primary data sources, though many of the findings appear to stem from interviews. This makes reliability assessment less easy. Secondary data sources (reports, policies, etc.) are used adequately throughout the report.
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?	No	As mentioned in assessment criteria 2.9, there is no discussion of potential bias or gaps in sources/data. The minimal number of focus groups with beneficiaries (only 1 discussion with 2 different beneficiary groups) in itself signals potential information gaps or bias.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The evaluators made a point of indicating that this was done.

4. Analysis and Findings	Yes No Partial	Assessment Level:	Fair
---------------------------------	----------------------	-------------------	-------------

To ensure sound analysis and credible findings

1. Are the findings substantiated by evidence?	Partial	The findings clearly stem from the data, but specific sources are not provided for qualitative data. As a result, the connection between the data and the findings is not always clear/varies. The basis for the findings is often unclear in the text, although many findings seem to be based on interviews.
2. Is the basis for interpretations carefully described?	Yes	Interpretations are carefully described.
3. Is the analysis presented against the evaluation questions?	Yes	Yes, the findings and analysis are structured by evaluation questions.
4. Is the analysis transparent about the sources and quality of data?	Partial	See sub-criteria 4.1, the analysis is inconsistent in its references to primary data from interviews, though the specific sources of (and other types of) qualitative data are occasionally referenced.

<p>5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?</p>	<p>Partial</p>	<p>The discussion of each output was prefaced by a brief discussion of the activities that led to their achievement. The evaluation report sought to use triangulation in order to explain the links between outputs and outcomes. The best example of this is on page 58 in reference to the increased/more consistent drug supplies at clinics. However, clear causal links weren't always made between activities, outputs and outcomes. In some cases, the casual links are shown. In others they are not. A case is when training is described: it is not clear in some cases how UNFPA was connected, other than through producing documents that were used. There is almost no analysis of the extent to which the training was used, although in several cases it was noted that changes in government programs made the training somewhat irrelevant.</p>
<p>6. Does the analysis show different outcomes for different target groups, as relevant?</p>	<p>Yes</p>	<p>The analysis refers to outcomes at the institutional (e.g. government programs) and beneficiary level in terms of different target groups.</p>
<p>7. Is the analysis presented against contextual factors?</p>	<p>Yes</p>	<p>Yes, contextual factors are well-described and the analysis is presented against these.</p>
<p>8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?</p>	<p>Yes</p>	<p>The analysis seeks to incorporate gender analysis into the discussion of findings, as well as a discussion of equity and vulnerability as it relates to socioeconomic status.</p>

<p>5. Conclusions</p>	<p>Yes No Partial</p>	<p>Assessment Level:</p>	<p>Very good</p>
<p><i>To assess the validity of conclusions</i></p>			
<p>1. Do the conclusions flow clearly from the findings?</p>	<p>Yes</p>	<p>The conclusions are linked to the evaluation questions and flow from the findings.</p>	
<p>2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?</p>	<p>Yes</p>	<p>Yes, the conclusions are generally broader than the findings and suggest main factors explaining the results and demonstrate an understanding of the complexity of the UN system.</p>	
<p>3. Do the conclusions appear to convey the evaluators' unbiased judgment?</p>	<p>Yes</p>	<p>There does not appear to be any inherent bias in the conclusions. Though brief, they were impartial and stemmed from the findings.</p>	

<p>6. Recommendations</p>	<p>Yes No Partial</p>	<p>Assessment Level:</p>	<p>Fair</p>
<p><i>To ensure the usefulness and clarity of recommendations</i></p>			
<p>1. Do recommendations flow logically from conclusions?</p>	<p>Yes</p>	<p>The recommendations are connected to the relevant conclusions in the presentation.</p>	

2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	They are clear and targeted, but there is no information on human, financial and technical implications in the majority of cases.
3. Do recommendations appear balanced and impartial?	Yes	There is no indication of partiality.
4. Is a timeframe for implementation proposed?	Partial	A specific timeline is not provided, however all recommendations are intended for the next program cycle.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	All of the recommendations are given high priority (and are thus not prioritised). However, they are clearly described so that a management response can be given.

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	There is a discussion of how gender is integrated into the analysis and indicators on page 19.
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?	2	The criteria are standard OECD-DAC criteria, but the indicators and hypothesis of the evaluation questions are formulated in a way to capture whether or not GEEW has been integrated into the design of the support provided.
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	3	The evaluators used methods described on page 19 of the report that determined the extent of integration of gender into the implementation of the program and in the results achieved. The evaluation methods were careful to include gender at two levels: I) Cómo el UNFPA ha previsto contribuir a la reducción de las inequidades basadas en género en el marco de acción de los otros productos (SSR, salud materna y neonatal, dinámicas de población y educación para la sexualidad). II) Cómo el UNFPA ha previsto que el enfoque de género (siempre en función del objetivo estratégico del UNFPA), sea un eje articulador en el contexto de coordinación del Sistema ONU. This was reflected in the interviews and field visits that took gender into account, by ensuring that both females and males would be interviewed and brought into focus groups and that, as necessary, data were disaggregated by gender.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	The findings reflect a gender analysis, and the conclusions (#8) and recommendations incorporate the analysis and demonstrate an understanding of challenges related to gender in various interventions and geographies. However, only one of nine conclusions mentions gender and none of the recommendations do.

(*) 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 totaling 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	11	7	82	
Overall assessment level of evaluation report			Fair	
	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?

The fair rating is based on unevenness in the evaluation design and in its applications in terms of findings. While the weaknesses of the methods and findings need to be considered, the fact that the conclusions are strong and clearly expressed, suggests that the evaluation can be used to help design the next country program.

• What aspects to be cautious about?

To the extent that the findings are based on a limited selection of regions (only two) where beneficiaries were interviewed, the generality of some conclusions needs to be read with caution.

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:

