

Organizational unit: **UNFPA**

Year of report: **2017**

Title of evaluation report: **Independent Evaluation of UNFPA's Regional Intervention Action Plan 2014-2017**

Overall quality of report: **Good**

Date of assessment: **15 July 2017**

Overall comments: The evaluation report is an example of sound structure and clarity throughout, especially, in terms of structure of the Findings chapter. The evaluation framework is clearly described in the Methodology section. The findings are structured according to the program themes. Each theme is assessed against five evaluation criteria and relevant evaluation questions. Each finding is assigned a number and supported by evidence. The analysis is transparent about sources and quality of data. The cause-effect links between outputs and outcomes are rationally explained. Conclusions follow logically from the analysis and present the most important issues that UNFPA needs to consider in decision-making and next cycle program development. Recommendations are very clear and action-oriented. GEEW is integrated in the evaluation scope of analysis, 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 <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:	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	Comment: 1. The evaluation report is presented in an easy to read and understand manner. The annexes include the required items. 2. The report is of 93 pages, excluding annexes. 3. Lessons learned are incorporated into the analysis, but they are not highlighted. We can easily find them in the text if we use Find-on-page tool. In the meantime, the annex 7 has a component "What Lessons Does EECA Have to Share?" among Focus Group Discussion Questions. 4. The annexes include Sampling Framework, Interview Framework, UNFPA EECA Region Staff Survey and Focus Group Discussion Questions. The annexes do not contain an evaluation matrix and one is not presented or described in the text of the report. The absence of evaluation matrix is a real concern, not only with regard to criterion 1, but, more importantly, for the ability to find the substantiating evidence of the evaluation findings, under criterion 5. 5. An executive summary has clear structure and functions as a stand-alone section. 6. The Executive Summary covers the required items and establishes a firm basis for understanding the evaluation conclusions and recommendations which are clearly explained. 7. The Executive Summary is concise, being 6 pages, one page longer than desirable.	
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		
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		
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		
<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		
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		
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes		

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	Comment: 1. While the Methodology does not describe the target audience, the target audience is clear " The Evaluation was commissioned by the EECA Regional Office to inform decision-making and next cycle programme development as per the Biennial Budgeted Evaluation Plan 2015-2016" (p. 1). 2. The regional context of the evaluation is thoroughly described (pp. 7-16). Explanation of general constraints was focused on methodological limitations (p. 20). 3. While there is no evaluation matrix that would describe the theory of change, it takes the Regional	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes		

<p>3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?</p>	<p>Partial</p>	<p>Plan as the theory, which is reasonable. The evaluation report describes the outcomes of the UNFPA regional program action plan (pp. 13-16) and comments on the theory of change (including assumptions) in the text where relevant. In describing the theory of change, the evaluators note that this had changed in 2016 in the light of reduction in resources, so the evaluators had reconstructed the logic. Given this, the theory of change used in the evaluation is reasonable.</p>
<p><i>To ensure a rigorous design and methodology</i></p>		
<p>4. Is the evaluation framework clearly described in the text and in the evaluation matrix?</p>	<p>Partial</p>	<p>4. The evaluation framework is clearly described in the Methodology section: purpose and objectives of the evaluation, scope, methods, criteria, tools, and analysis, stakeholders' interviews, limitations, and considerations. However, there is no evaluation matrix in the report. Methodology establishes the evaluation questions and explains methods for data collection. Data sources are explained in the annexes. Chapter 3 "Evaluation findings" contains tables with program outputs and outputs' indicators. Outcomes are presented in the Table 11 "EECARO Outcome Report 2014-2016" of the Conclusion section.</p>
<p>5. Are the tools for data collection described and their choice justified?</p>	<p>Yes</p>	<p>5. The tools for data collection are described on the p. 19. 6. Sampling Framework in the annex 2 could be considered as a stakeholder map. The stakeholder consultation process is briefly described on the p. 19 which mentions the stakeholder mapping exercise, but it does not include description of the consultations with key stakeholders on draft recommendations.</p>
<p>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)?</p>	<p>Yes</p>	<p>7. Actual section "Evaluation tools and analysis" does not describe analysis methods. It is briefly noted in the limitations section (2.3.6.) that "it proved difficult to judge the level of programme understanding of the respondents and their personal views due to the high percentage of "don't know" and "undecided" responses..." (p. 20).</p>
<p>7. Are the methods for analysis clearly described for all types of data?</p>	<p>Yes</p>	<p>8. Methodological limitations and mitigation measures are explained clearly "Limitations were mitigated to the extent possible through triangulation of data and with support from UNFPA staff..." (p. 20). 9. There is indication of a sampling strategy being developed and applied in the text (pp. 19-20) and the annex 2. The survey is a 100% sample. While annex 5 does not show clearly how the sampling framework was applied (reference to the annex on the p. 20), the limitations section indicates that there was a relatively low rate of response.</p>
<p>8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)</p>	<p>Yes</p>	<p>10/11. The methodology appears to enable the collection/analysis of disaggregated data and be appropriate for assessing cross-cutting issues. It is said on the p. 18 that "The Evaluation was designed to be transparent, inclusive, participatory, as well as gender- and human rights responsive." Also, the evaluators mention that "attention to gender equality and human rights was paid in developing the interview tools" (p. 19).</p>
<p>9. Is the sampling strategy described?</p>	<p>Yes</p>	
<p>10. Does the methodology enable the collection and analysis of disaggregated data?</p>	<p>Yes</p>	
<p>11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?</p>	<p>Yes</p>	

<p>3. Reliability of Data</p>	<p>Yes No Partial</p>	<p>Assessment Level:</p>	<p>Good</p>
<p><i>To ensure quality of data and robust data collection processes</i></p>			<p>Comment:</p>
<p>1. Did the evaluation triangulate data collected as appropriate?</p>	<p>Yes</p>	<p>1. There is clear evidence of the triangulation. The data were collected from different sources: documents review, interviews, a largely quantitative online survey, thematic discussion groups. The report notes that "Unfortunately, at two days each per country, the Team were unable to collect in-depth data although very useful information was gleaned which complemented the key respondents' information and provided triangulation for the data already collected" (p. 20). 2. The report has annexes which include detailed information about qualitative and quantitative data sources (Sampling Framework, List of Persons Interviewed, Background Documents Reviewed, Interview Framework, UNFPA EECA Region Staff Survey, Focus Group Discussion Questions).</p>	
<p>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</p>	<p>Yes</p>	<p>3. Information about limitations and mitigation measures is presented on the p. 20. 4. Section 2.3.7 "Ethical Considerations, Conflicts of Interest" contains information about the issues of discrimination and other ethical considerations (p. 20). Also, we can see from the annexes that the evaluators addressed these issues during the data collection. For example, it is said in the annex 6 "The survey is completely anonymous" (p. 105). In the meantime, the annex 7 "Focus Group Discussion Questions" does not explain how the process was organized in terms of sensitivity to issues of discrimination and other ethical considerations.</p>	
<p>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?</p>	<p>Yes</p>		
<p>4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</p>	<p>Partial</p>		

<p>4. Analysis and Findings</p>	<p>Yes No Partial</p>	<p>Assessment Level:</p>	<p>Good</p>
<p><i>To ensure sound analysis and credible findings</i></p>			<p>Comment: This report could be an excellent example of organizing the findings. The structure of the</p>

1. Are the findings substantiated by evidence?	Yes	chapter helps to present the analysis and interpret the data in a consistent way. 1. Findings are supported by evidence (documents reviewed, discussion with key informants, and others). They are structured according to the regional plan (as revised in 2016) which compensates for a lack of an evaluation matrix. The evaluators collected enough data about the volume of outputs produced by the program. Outputs, indicators, and targets met are presented in the tables across the report. These, along with interviews, also provide indicators of outcomes, to the extent that these have been measured.
2. Is the basis for interpretations carefully described?	Yes	2. Interpretations are based on comparison of data collected from different sources - an example is the statement that "There are many other examples of where the EHP has been able to influence country-level work..." (p. 32). 3. The findings are structured according to the program themes, which compensate for the missing evaluation matrix. Each theme is assessed against five evaluation criteria and relevant evaluation questions. Each finding is assigned a number. An evaluation matrix which would show the evaluation questions and how they were to be measured, would have made the evaluation questions clearer.
3. Is the analysis presented against the evaluation questions?	Partial	4. The analysis is transparent about sources and quality of data - often citing, for instance "key informants... are of the opinion that..." "some staff ... mentioned that..." "the evaluation team feels that..." (p. 32), and "this mistaken belief is due to..." "The development of the SWIT tool is described in Box 1" (p. 33).
4. Is the analysis transparent about the sources and quality of data?	Yes	5. The cause-effect links between outputs and outcomes are rationally explained in the analysis of findings, for instance: outputs "RP's work on gender disaggregation of data, as well as support for censuses and improving systems for vital registration" were essential to the outcome "quality of public policy and planning in the region" (p. 57). Another example is the statement that "Where PD has contributed to improvements in quality of census data and other regular surveys, this assumption may very well be valid. Important examples exist of advocacy products (particularly policy briefs) that have utilised improved PD methodologies and approaches" (p. 59). The evaluators also conclude that "However, accountability for outcomes is weak, both in terms of the component's contribution to the SP goal, as well as in understanding if capacity fostered through training and other capacity building interventions is sustained..." (p. 60).
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Even though all the tables discuss the progress on outputs' indicators, the evaluators explain the connection between outputs and outcomes in the narrative text. Ideally, effectiveness is about the effects/changes within society (outcomes) due to the outputs produced under the regional plan. Thus, the analysis needs to prioritize the outcomes (outcomes indicators) vs outputs (outputs indicators). 6. The analysis shows different outputs and outcomes for different countries and beneficiaries of the UNFPA programs like young male refugees, boys and girls, policymakers, and national partners (pp. 23-24).
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	7. The report presents the analysis against contextual factors which is evident from the text, for instance, the current RIAP "challenges the accepted status quo of traditional family composition and values..." (p. 34). Also, the evaluators explain that "the Evaluation considered the contextual factors that may have affected the implementation of RP interventions and their potential to bring about desired outcomes. Where regional outcome-level data were lacking, the Evaluators were limited in assessing the extent to which interventions have contributed to the achievement of results" (p. 18).
7. Is the analysis presented against contextual factors?	Yes	8. The analysis elaborates on cross-cutting issues in all sections of the report: Sexual and reproductive health (adolescents and young people – p. 23), HIV (young people, marginalized populations, inclusivity, inequality, universality and community empowerment – p. 34), Humanitarian (adolescents and youth, gender), Adolescents and youth (capacity development of young people to engage in policy advocacy – p. 46), Gender (human rights, GBV, male involvement and harmful practices like child marriage and FGC (p. 49), Population and development (women and youth), and Cross-cutting issues (p. 63).
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	

5. Conclusions	Yes No Partial	Assessment Level:	Fair
<i>To assess the validity of conclusions</i>		Comment:	
1. Do the conclusions flow clearly from the findings?	Partial	1. Conclusions are related to the evaluation questions which separate the chapter into the sub-sections. This is not in conformity with the UNFPA Handbook and the connections with findings are not always clear. While the conclusions are logical and easy to understand, they are presented as narrative text and are not numbered. 2. The conclusions present the most important issues that UNFPA needs to consider in decision-making and next cycle program development and go beyond the findings. For instance, the conclusions	

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	state that "...RIAP planning processes could be better aligned to favour country needs and priorities and consult more effectively in the development of the RIAP" (p. 75), "...indicators are not consistently well aligned with actual programming priorities and targets not always meaningful..." (p. 77), "a second important constraining factor is the growing political conservatism in the region" (p. 77). There are references to the sources of data like the results tables (p. 77), interviews, and surveys (p. 80). The evaluators discuss the findings and draw generalizations from them that are broader than the findings. For example, they conclude that "The RP has met with several important challenges" (p. 77), "A number of the RP programme outcomes align with the philosophy of the 2030 Agenda... However, this work is considered both uneven across the regional programmed, and would benefit from greater strategic framing and cross-programme collaboration" (p. 80). Table 11 provides EECARO Outcome Report (2014-2016) which is clear and short. 3. The conclusions convey an unbiased judgment.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	

6. Recommendations	Yes No Partial	Assessment Level:	Fair
<i>To ensure the usefulness and clarity of recommendations</i>		Comment:	
1. Do recommendations flow logically from conclusions?	Yes	1. Recommendations flow logically from conclusions and are presented after them in the same chapter. This is not a common approach in UNFPA and it may confuse conclusions with recommendations.	
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. The recommendations are specific and action-oriented and, while the evaluators do not mention intended users directly, it is obvious that they mean the UNFPA. However, it is not clear whether they are referring to UNFPA Headquarters or regional offices and this makes implementation monitoring more complex. The evaluators are clear about what and how should be done, for instance, they recommend to continue "consultative annual planning; the regional newsletter..." based on "advocacy and communications..." (p. 76), "The strategy should include specific attention and identification of resources for training (and other CD intervention) follow up..." (p. 80).	
3. Do recommendations appear balanced and impartial?	Yes	3. Recommendations look balanced and impartial but there is no evidence of consultation before the recommendations were drafted.	
4. Is a timeframe for implementation proposed?	Yes	4. Even though there are a total of 15 recommendations, they are all valuable in terms of measures to improve the UNFPA regional plan. The timeframe is "decision-making and next cycle programme development," which suggests the next programme formulation cycle.	
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	5. The recommendations are prioritized (Medium to High). There are just 1-2 sentences per each recommendation, and the formulations are SMART to facilitate appropriate management response and follow up on each specific recommendation.	

7. Gender	0 1 2 3	Assessment Level:	Good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>		Comment:	
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>1. Scope and indicators. Scope of the evaluation directly reflect GEEW as a program component: (1) SRH, including HIV and Humanitarian; (2) Gender Equality, (3) A&Y; and (4) PD. Also, gender is mainstreamed into all program areas. There is a GEEW-responsive objective "To determine the extent to which UNFPA's regional programme takes into consideration cross-cutting issues such as inequality and human rights based approaches, and gender equality in programme design and implementation" (p. 87). Results tables include GEEW dimensions into the indicators, for instance: "(ii) GBV referral to SRH services and provision of comprehensive SRH services to GBV survivors..." (p. 24), "child and forced marriage" (p. 53), "Countries with gender equality national action plans that integrate reproductive rights..." (p. 79). In some cases, the evaluators do not disaggregate marginalized adolescents, young people, disadvantaged and vulnerable populations, but in other cases they do disaggregate (adolescent girls and youth).</p> <p>3 We can see from the report that the evaluators collected sufficient information to measure progress on GEEW: "...on implementation of the Global Standards on Essential Services for adolescent girls and young women..." (p. 50), "The Guidance Note covers the role of gender equality within the 2030 Agenda for Sustainable Development..." (p. 54), "The migration of women and girls remains largely unregulated... levels of GBV/violence against women are high" (p. 55).</p> <p>2. Evaluation criteria and evaluation questions. There is a criterion Effectiveness which includes GEEW-related evaluation question: "To what extent has the human rights-based approach been implemented in RIAP supported interventions? (p. 88). There is a specific evaluation question EQ 2E which addresses GEEW: "To what extent did the RIAP utilise a human rights-based approach and incorporate principles of gender equity in programme design and implementation?" (p. 58).</p> <p>3. Methodology.</p>	

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 authors of the report mention that "Limitations were mitigated to the extent possible through triangulation of data" (p. 20). Annex 3 "List of Persons Interviewed" does not disaggregate people according to sex. The evaluators say in the methodology that "attention to gender equality and human rights was paid in developing the interview tools" (p. 19). The authors do not explain how they addressed gender equality and human rights issues during the design and implementation of the evaluation. Also, the evaluators do not provide any specific approaches to make visible diverse perspectives and promote participation in the evaluation of both women and men from different stakeholder groups. The report does not explain how evaluation data collection tools were gender-responsive.
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	The authors do not describe how ethical standards were applied throughout the evaluation, even though they mentioned that "The Evaluation followed the guidance on the integration of gender equality..." (p. 18). We can find a note in the Annex 6 of the UNFPA EECA Region Staff Survey that "The survey is completely anonymous" (p. 105). 4. Findings, conclusions, and recommendations. Evaluation findings reflect a gender analysis: the chapter 3.5. is devoted to the Gender program area, gender is mainstreamed into the other parts of the analysis as well: 3.1. Sexual and Reproductive Health, 3.2. HIV, 3.3. Humanitarian. But, we can find different examples in the chapter 3: sometimes the evaluators mention "young adolescent girls" (pp. 44, 45, and 47), more often they refer their data to general target groups like Adolescents, Young people, Vulnerable and Marginalized Youth (pp. 47, 48). 4.2. In general, the evaluation conclusions reflect a gender analysis, for instance, women and girls are mentioned on the p. 80. There is conclusion on EQ 2E "To what extent did the RIAP utilize a human rights-based approach and incorporate principles of gender equity in programme design and implementation?" (p. 80). 4.3. Evaluation recommendations do not include any GEEW-specific recommendations. The evaluators missed opportunity to disaggregate target groups in terms of gender in the section 4.2.2 "Effectiveness Recommendations": "Develop, a specific cross-programme strategy to address the needs of marginalised populations, taking into consideration the RP and CP experience with Y/KPs, Roma, migrants and refugees, including IDPs; and embed the strategy in the 2030 Agenda including mapping to key approaches and expected outcomes" (p. 80).
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	

(*) 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		65	35	
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

Within some limitations in terms of time and the size of the evaluation team, the evaluation is very consistent with UNFPA norms and, as the comments show, provides a very good assessment of the regional programme.

Consideration of significant constraints

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

Yes

No

If yes, please explain: