



Organizational unit: **Iraq Country Programme** Year of report: **2021**

Title of evaluation report: **United Nations Population Fund (UNFPA) Evaluation of the 2nd Country Programme 2016 - 2019**

Overall quality of report: **Very good** Date of assessment: **20 July 2021**

Overall comments: This is a thorough evaluation of the Iraq 2nd Country Programme 2016 – 2019 and covers activities implemented and results achieved in all the thematic areas of programming (sexual and reproductive health, adolescents and youth, gender equality and women’s empowerment, and population dynamics) being assessed against OECD-DAC criteria: relevance, effectiveness, efficiency and sustainability in addition to assessing connectedness and coordination of the programme as a whole. It covers the period from January 2016 to December 2018 and intends to enhance the accountability of the UNFPA Iraq CO for the relevance and performance of its country programme and broaden the evidence base for the design of the next Country Programme (CP) cycle. The executive summary serves as a stand-alone document although the recommendations in this section could be more clearly presented if numbered and separated into strategic and programmatic recommendations. The methodology section of the report could be improved by describing how ethical considerations were addressed in the evaluation processes. The findings are thorough. Both the conclusions and recommendations sections of the evaluation are well formulated and presented. Gender is integrated throughout the findings, conclusions, and recommendations; and the evaluators also considered the extent to which the program reached people with disabilities.

Assessment Levels

- Very Good** (blue box): strong, above average, best practice
- Good** (green box): satisfactory, respectable
- Fair** (yellow box): with some weaknesses, still acceptable
- Unsatisfactory** (red box): 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:	Good
--	----------------------	-------------------	-------------

To ensure the report is comprehensive and user-friendly

1. Is the report structured in a logical way? 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? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Yes	The report is clearly structured, easy to navigate, and well written.
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)	Partial	At 72 pages (excluding the Annexes) the report exceeds the limit by 2 pages. It is text heavy in some places, and some of the presentation could be more succinct to increase readability (such as the Findings Summary for Efficiency on p 54). To help meet the page limit, in addition to being more succinct, some of the more detailed information could have been allocated the annexes.
3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?	Yes	All required elements are included in the Annexes.

Executive summary

4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) intended audience; iv) Methodology; v) Main results; Vi) Conclusions and Recommendations?	Partial	The executive summary serves as a stand-alone document, however, the recommendations could be more clearly presented if numbered and separated into strategic and programmatic recommendations rather than being consolidated into paragraphs.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	At around 4 pages, the summary is reasonably concise.

2. Design and Methodology	Yes No Partial	Assessment Level:	Very good
----------------------------------	----------------------	-------------------	------------------

To ensure that the evaluation is put within its context

1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The context and background of the intervention and evaluation are comprehensively described including development challenges, national strategies, and humanitarian response including statistics on population dynamics, SRH, and GEEV.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Partial	Evaluators noted that they tested the intervention logic underlying the CP to assess the extent to which the expected results of the CP have been achieved. They did not revise it even though the stated outputs are actually outcomes and the stated outcomes are the titles of the 4 program areas. The ToRs indicated that the ToC should be reconstructed.

To ensure a rigorous design and methodology

3. 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 matrix is described in the text, and attached as Annex 6. The latter includes detailed list of information gathered on the evaluation questions, assumptions to be assessed, Indicators, sources of information, and methods and tools for data collection.
4. Are the tools for data collection described and their choice justified?	Yes	The evaluators clearly describe data collection tools and how they were used. The benefits of using the complementary methods is explained.
5. 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	There is a specific subsection on stakeholder participation. It explains the range of stakeholders consulted, including women and youth from camps, and notes that an Evaluation Reference Group was formed. A validation workshop was also held with key stakeholders on findings and recommendations.
6. Are the methods for analysis clearly described for all types of data?	Partial	It is noted that data analysis was structured according to the evaluation criteria and included an assessment of what answers the data suggested for each of the evaluation questions. However, the specific process/mechanics of data analysis methods could have been more clearly presented.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	Limitations relating to data collection and security are described along with their mitigation strategies.
8. Is the sampling strategy described?	Yes	Evaluators note that based on the evaluation purpose, a purposive sampling approach for the selection of stakeholders to participate in KIIs and FGDs. The criteria for the selection of implementing partners and UN agencies is included.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The report includes disaggregated data, and the list of persons consulted is disaggregated by stakeholder type and gender. The analysis looks at the extent to which the CP was disability inclusive but does not provide specific data in this regard.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	Evaluators assessed the integration of human rights and gender equality in the 2nd UNFPA Iraq CP (2016-2019) as part of the effectiveness criterion.
3. Reliability of Data		
	Yes No Partial	Assessment Level: Very good
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	Evaluators adopted a participatory approach to include a diverse range of governmental and non-governmental stakeholders and have used mixed methods to allow triangulation.□
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation describes various sources of data used, both qualitative and quantitative. In addition, challenges related to absence of reliable data due to context are also explained.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	The evaluation report makes reference to UNEG Norms and Standards for Evaluation and the UNEG Ethical Guidelines for Evaluation, and the introductory section of the data collection protocols show that confidentiality was addressed. The evaluators could have gone further by more fully describing how ethical considerations were applied, particularly for focus groups.
4. Analysis and Findings		
	Yes No Partial	Assessment Level: Very good
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	Evaluation findings are backed up by qualitative as well as quantitative data. Sources of data are mentioned in footnotes.
2. Is the basis for interpretations carefully described?	Yes	Interpretations in the findings are clearly described. Terms commonly used include "assessment/report revealed ...". and "interviews with stakeholders and beneficiaries indicate ...".
3. Is the analysis presented against the evaluation questions?	Yes	Evaluation findings are structured according to the main criteria and start with the evaluation questions and include specific findings in summary form followed by detailed analysis.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Causal linkages between outputs and results are reported, wherever observed. For example, evaluators note that "Awareness-raising activities also resulted in some men accepting to attend awareness-raising sessions on various topics".
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings analyze and present differential results for target groups. For examples it is noted that "While programming focused on some marginalized groups, in particular former wives/widows and children of ISIL fighters in IDP and refugee camps, focus on marginalized groups was generally limited, including on people with disabilities and female-headed households".
6. Is the analysis presented against contextual factors?	Yes	Contextual factors, especially security and instability are analyzed including their impact on intervention.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Evaluation findings include analysis of cross-cutting issues such as gender equality and vulnerability. Disability inclusion is also briefly discussed, mainly in terms of its evolution within the CP since 2018.

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	Conclusions are clearly derived from findings and are followed by number of the respective evaluation questions contained in the text of the finding section. They are organized as strategic conclusions and programmatic 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 and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Yes	The conclusions are formulated more broadly than the findings and give a clear understanding of the intervention. Additionally, they cover appropriate cross-cutting issues including human rights and gender. □	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	Conclusions clearly stem from the findings and do not reflect any 'bias'.	
6. Recommendations	Yes No Partial	Assessment Level:	Very good
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	Recommendations are categorized by strategic and programmatic level, are drawn from the conclusions and indicate the number of conclusions on which they are based.	
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	Recommendations are directed towards intended users and are action oriented with details on operational implications.	
3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?	Yes	Recommendations are practical, useful, and cover relevant key cross cutting issues such as gender and vulnerability.	
4. Are the recommendations prioritized?	Yes	Recommendations are prioritized.	
7. Gender	0 1 2 3 (**)	Assessment Level:	Very good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>			
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	<p>a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3) GEEW is not explicitly highlighted in the objectives or scope of the evaluation. = 0</p> <p>b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3) The evaluation does not have a standalone criterion on gender or human rights, however, it was assessed as part of the effectiveness criterion. = 3</p> <p>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) EQ7 "To what extent did the implementation of the UNFPA Iraq CP (2016-2019) take into account gender equality and human rights principles?" focuses on gender equality and human rights. = 3</p> <p>d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results? (Score: 0-3) Evaluation matrix includes indicators on vulnerable population and gender equality. = 3</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	3	<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3) Evaluators note in the methodology that efforts were made to ensure that both women and men working for IPs of the UNFPA Iraq CO were interviewed and that the beneficiaries consulted included both women and men, and key populations including adolescents and youth, refugees, IDPs and vulnerable and marginalized groups. However it would have been useful to describe steps taken to ensure the evaluation process was gender responsive. = 2</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) The evaluation adopted a mixed-methods approach to collect data which is appropriate to evaluate GEEW considerations. = 3</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Sources are diverse as data was collected from a wide range of stakeholder groups which included focus groups with rightsholders. = 3</p> <p>d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3) Rightsholders consulted by the evaluation team include refugees, IDPs and vulnerable and marginalized groups. =3</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) The evaluation makes reference to UNEG Norms and Standards for Evaluation and the UNEG Ethical Guidelines for Evaluation. In addition, it is noted that considering confidentiality issues, the IPs managing women's centres were contacted to identify women beneficiaries who would be willing to participate in an FGD held by the evaluation team. = 3</p>	

<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3) Background has a dedicated section on gender equality and women's empowerment (Section 2.6) and covers issue related to human rights and gender equality including FGM, child marriage, and GBV. = 3</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3) Triangulation is observed throughout the evaluation findings along with use of quotes from different stakeholder groups, including GBV survivors, however, it would have been helpful to highlight voices of different groups a bit more. = 2</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) These are not described. = 0</p> <p>d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3) GEEW issues are adequately covered in the recommendations including GBV, FP, and SRH. = 3</p>
--	---

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted.

(**) Scoring uses a four point scale (0-3).

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	93	7		
<i>Overall assessment level of evaluation report</i>	Very good			