

Organizational unit:	Palestine CO	Year of report:	2022
Title of evaluation report:	EVALUATION OF THE UNITED NATIONS POPULATION FUND (UNFPA) PALESTINE 6th COUNTRY PROGRAMME (2018 – 2022)		
Overall quality of report:	Good	Date of assessment:	23 June 2022
Overall comments:	<p>This evaluation provides a solid assessment of the UNFPA State of Palestine 6th Country Programme (CP) (2018-2022). The methodology included document review, key informant interviews and focus group discussions with key stakeholders, implementing partners and beneficiaries. Notably, there was a total of 245 participants. Although the evaluation participants are disaggregated by stakeholder group and gender, the listing of interviewees is not included in Annexes. The evaluation matrix is comprehensive and includes indicators and summarized key findings for each evaluation question. A shortcoming of the evaluation is that findings could be more consistently substantiated with specific sources of evidence, particularly with secondary/documentary sources. As well, several of the recommendations are high level and could be more specific and actionable. The report could also have benefited from final proofreading. Nevertheless, the evaluation performs strongly in respect to GEEW and disability inclusion. Disability is addressed as one of the cross-cutting themes and within three evaluations questions (under Relevance, Effectiveness and Coverage). Overall, the evaluation appropriately engages the most-at-risk populations (MARPs) including people with disabilities, and gender-based violence survivors.</p>		
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 assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)</i>		
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 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)?	Partial	The report is well structured, following the expected logic. However, the report would have benefited from final proofreading as it includes a number of spelling and punctuation errors, mixed font in some paragraphs, and some text from earlier versions that was not updated. It would also benefit from more visual aids to convey key information. Notably, right-based terminology is used throughout.	
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 reasonable in length with 67 pages. Additionally it includes annexes, leading to a report of 220 pages.	
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)?	Partial	The annexes include the ToRs, list of documents consulted, evaluation matrix, data collection tools, reconstructed ToC, among others. Annexes do not include however a list of stakeholders interviewed (although Table 2 shows aggregated number of individuals interviewed by type of stakeholder, i.e. government, UN, local NGO, etc.).	
<i>Executive summary</i>			
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?	Yes	The executive summary includes all of the essential elements: main purposes, objective, a brief description of Palestine's 6th Country Programme, main audience and primary users, methodology, key findings overall by criteria, main conclusions and recommendations.	
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is concise at 5 pages.	
2. Design and Methodology	Yes No Partial	Assessment Level:	Very good
<i>To ensure that the evaluation is put within its context</i>			
1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The country context is clearly described, including development challenges and national strategies, SRHR, the situation of adolescents and youth and GEEW.	
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Partial	Revisions to the ToCs for each programme area are referred to under section 3.2.3, although the text was not updated from an earlier version of the report as it states "The ToC that describes how and why . . . will be provided in the next version of this draft report". Reconstructed versions are shown in Figure 7 and in Annex 5. Nevertheless, the reconstruction involved mostly adding end dates to outputs and did not address some structural issues such as measurability of the results (i.e., outcome-level results being framed as 'enhanced capacity') and some mixing of outputs and outcomes (i.e., "national youth strategy is operationalized by the end of 2022" is presented as an output).	
<i>To ensure a rigorous design and methodology</i>			

<p>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?</p>	Yes	<p>The evaluation matrix is organized by criteria and includes evaluation questions, assumptions, indicators, sources of information, and methods and tools for data collection. The description under sources of information is detailed, providing good guidance of where secondary data should be coming from. Each criteria in the evaluation matrix has key findings.</p>
<p>4. Are the tools for data collection described and their choice justified?</p>	Yes	<p>The evaluation utilized several data collection methods, including key informant interviews (KIIs), FGD, and desk review. They are all briefly described and the rationale for their selection is explained.</p>
<p>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)?</p>	Yes	<p>The report includes a comprehensive stakeholders' map (Annex 4). The stakeholder consultation with diverse actors and beneficiaries is also described in detailed. According to the report, "the recommendations will be fine-tuned in a consultative process, because of participatory discussion with CO Palestine and follow-up rounds of validation with the Evaluation Reference Group"; since the sentence is written in future tense it is not clear whether the evaluation team did undertake the consultation process or not, however there was clear intent to do so. The stakeholder mapping and data collection methods involved women, adolescents, youth, men, most-at-risk populations (MARPs) including people with disabilities, and gender-based violence survivors.</p>
<p>6. Are the methods for analysis clearly described for all types of data?</p>	Yes	<p>The evaluation mainly used qualitative methods for data collection including document review, interviews, focus group discussions and observations. Quantitative data was compiled from secondary sources from desk review.</p>
<p>7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?</p>	Yes	<p>Limitations of the evaluation are well described as well as mitigation strategies to minimize them. The major limitations noted were the use of remote data collection methods, particularly for those participants with restricted internet and phone access, and the challenges of attribution, particularly for the SRHR and gender equality programming.</p>
<p>8. Is the sampling strategy described?</p>	Yes	<p>The sampling of stakeholders who participated in the KIIs and FGDs was based on a purposive and participatory approach. The sampling strategy took into consideration the gender and diversity factors and vulnerability guided by the UNFPA Evaluation Handbook. The universe, the report mentions, was identified based on the stakeholders map provided by "UNFPA Palestine", the initial review of programme documents and discussions with the UNFPA team during the design phase. A total of 245 people participated; 81 duty bearers and 164 beneficiaries, as presented in Table 2.</p>
<p>9. Does the methodology enable the collection and analysis of disaggregated data?</p>	Partial	<p>The selected methodology allowed for disaggregation by gender (162 females and 83 males). However the disaggregated data on beneficiaries with disabilities and gender-based violence survivors is not presented even though the FGDs were structured so that this information could be collected since they were convened per type of intervention and beneficiary group.</p>
<p>10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?</p>	Yes	<p>Both the design and chosen methodology were appropriate for assessing cross-cutting issues. EQ1, EQ2 and EQ9 address MARPs. FGDs and KII ensured participation of vulnerable groups in the data collection, such as people with disabilities and gender-based violence survivors.</p>
<p>3. Reliability of Data</p>	<p>Yes No Partial</p>	<p>Assessment Level: Good</p>
<p><i>To ensure quality of data and robust data collection processes</i></p>		
<p>1. Did the evaluation triangulate data collected as appropriate?</p>	Yes	<p>The evaluation matrix includes diverse data sources, appropriate for triangulation. The evaluation matrix shows clear intent to collect triangulated data for each evaluation question from multiple sources and stakeholders with different perspectives.</p>
<p>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</p>	Partial	<p>Qualitative and quantitative data sources are identified in general terms under the methodology section. However, since there is no Annex with interviewed stakeholders/implementing partners listed, it is not clear from the stakeholder map provided who participated in the evaluation. It is only clear that 245 people were involved in total, and from which stakeholder group (table 2 and figure 1). Reliability is discussed under section 1.3.6 Data Validation and Analysis.</p>
<p>3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</p>	Yes	<p>Ethical considerations are clear regarding sensitivity to specific issues. Oral consent was obtained from all participants who took part in this evaluation. According to evaluators, special needs around GBV, and disability-related work were considered, while ensuring confidentiality with adequate and informed consent. In addition, to ensure privacy, FGDs involving female participants were organised separately from those for male participants, particularly for SRH and GBV to ensure that female and male participants felt comfortable, open, and uninfluenced by power dynamics when discussing sensitive issues such as reproductive health care practices.</p>
<p>4. Analysis and Findings</p>	<p>Yes No Partial</p>	<p>Assessment Level: Good</p>

<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Partial	The findings could be more consistently substantiated with specific sources of evidence (cited interviews, documents, etc.). Although this is done well in some sections such as Relevance, in other cases sources are described more generally - quite frequently "a study/assessment" is mentioned but the document is not identified and in some places quotes from beneficiaries/stakeholders are included but not further substantiated with desk review. For example, in the discussion on Addressing chronic maternal health gaps (p 29) highlights that there is an "information rich maternal mortality report on the year 2020 that has come out recently" but the title of the report is not provided, and another study is mentioned in the discussion on family planning (p 31) but not identified.	
2. Is the basis for interpretations carefully described?	Yes	Findings are supported by a thorough analysis of the evidence. For all criteria there is a good explanation of each main findings with examples regularly being provided.	
3. Is the analysis presented against the evaluation questions?	Yes	The evaluation questions are listed for each criteria, and are used to structure the analysis.	
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Causal effects are shown. This is particularly clear under effectiveness where tables are used to show performance assessments against CPD indicators. Under the same criterion, a subsection on unintended effects described several negative and positive outcomes, i.e., in the area of GEWE, "while the focus of services was on adult married women, the benefits went beyond women themselves (many of the women participating in the FGDs reported that they gained awareness, skills and overall empowerment was felt by all other family members)."	
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The findings clearly discuss the extent to which the CPE has reached vulnerable groups, including young men and male and female adolescents, and those with disabilities. It also notes the additional vulnerabilities of those living in Gaza and in the more remote communities in the West Bank.	
6. Is the analysis presented against contextual factors?	Yes	Findings present a clear analysis of contextual factors, especially the political context and COVID-19.	
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The analysis elaborates on cross-cutting issues, including noting where the needs of specific groups are not sufficiently addressed (LGBTQ groups, persons with disabilities, etc.).	
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 findings. The evaluators have specified the respective evaluation questions, evaluation criteria and associated recommendation 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 and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Yes	The conclusions reflect the findings and are organized by two categories: strategic and programmatic level. They also reflect cross-cutting issues, for instance conclusion 4 describes weaknesses such as in the integration of LGBTQI populations and persons with disabilities in UNFPA programmes.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions appear to show the evaluators' unbiased judgement by presenting balanced and independent arguments.	
6. Recommendations			
	Yes No Partial	Assessment Level:	Good
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	Recommendations are appropriately drawn from conclusions, and specify the conclusion 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)?	Partial	Despite recommendations include implications, several are at a high level and do not appear to be specific enough to be actionable. For example, recommendation 4 mentions "The technical implication is that UNFPA should provide technical support to national and governorate governments to facilitate accountability mechanisms for the implementation of policies and laws [related to the ICPD agenda] ."	
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	The recommendations appear balanced. There is no evidence of bias. They also address cross cutting issues, such as "coordination mechanisms on young people, SRH and GBV."	
4. Are the recommendations prioritized?	Yes	Recommendations are presented as either high or medium priority.	
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?	<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) As per the ToR, the evaluation covered cross-cutting issues of population and development, human rights and gender equality, disability (3).</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 include a standalone criterion on gender and human rights specifically, but is mainstreamed into other criteria. The coverage criteria addresses vulnerable population particularly. (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). Under relevance, effectiveness and coverage, EQ 1, 3 and 8 cover vulnerable groups and in particular women and with specific focus on GBV survivors. (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) EQ 4, under effectiveness, assesses the monitoring of gender and human rights perspectives. (3)</p>
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	<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) The methodology notes how gender and inclusion were addressed, and specifically how interviews considered ethical safeguards when collecting information on GBV. (3)</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) Although more based on qualitative methods, it is appropriate for evaluating GEEW considerations within CPE context. (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) The evaluation consulted a wide range of data sources. However, some findings as noted above do not adequately triangulate data. (2)</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) The evaluation reached a diverse group of stakeholders (including beneficiaries such as women, adolescents, youth, men, most-at-risk populations, i.e. people with disabilities, and gender-based violence survivors). However, a list of stakeholders is not included. (2)</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) Ethical considerations are clearly explained as to special needs around GBV, and disability-related work were considered, while ensuring confidentiality with adequate and informed consent. Evaluators also ensured privacy in FGDs involving female participants which were organised separately from those for male participants. (3)</p>
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	<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) The contexts includes subsections on specific social groups and normative instruments to address human rights and gender equality are mentioned. (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) Findings include an adequate gender analysis and presents disaggregated data where possible. (3)</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Unintended effects are mentioned, both negative and positive - an example of the former being with the process of receiving cash payments. (3)</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) Recommendations include GEEW issues, such as Recommendation 10 which suggests more work can be done on protection and empowerment of women. (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	31	69		
Overall assessment level of evaluation report	Good			

The evaluation integrates adequately cross cutting issues of gender equality, human rights and disability inclusion, even though is not included as part of the evaluation objective.

Yes

Consideration of significant constraints (e.g. COVID-19 or civil unrest)

The quality of this evaluation report has been hampered by exceptionally difficult circumstances: Yes No

If yes, please explain:

COVID-19 mostly affected the data collection phase, and as such some data had to be collected remotely and therefore depended on respondents having access to Internet and telephones enabling remote communication.