

**Organizational unit:** Ukraine **Year of report:** 2022

**Title of evaluation report:** Evaluation of the 3rd UNFPA Country Programme for Ukraine (2018-2022): Final Evaluation Report

**Overall quality of report:** **Very good** **Date of assessment:** 21 December 2022

**Overall comments:** This Evaluation of the 3rd UNFPA Country Programme for Ukraine (2018-2022) is thorough and well-written. The complexity of the country context is clearly explained, noting political, security and economic challenges as well as population dynamics; however, the possible impact of the context on the methodological approach and sampling is recommended to be elaborated on, especially considering very few end beneficiaries were consulted. Though the evaluation report notes that other stakeholders were interviewed to fill this gap and speak on behalf of beneficiaries, it does not indicate which stakeholders specifically were called on to speak on behalf of vulnerable groups (for example, more community-based organizations representing persons with disabilities). Despite this, the evaluation did utilize a mixed-methods approach and presented a balance of data across sources in order to validate findings, including presenting disaggregated data by gender, age, disability status (as available and relevant). Overall, 86 KIIs were conducted across government, NGO, donors and UN agencies and 35 survey responses were received. Although the survey response rate is quite low (28%), the evaluators demonstrated how the respondents covered all stakeholder types and UNFPA programme component areas. The analysis is presented against the evaluation questions, with summary findings for each question presented prior to the more in-depth analysis. Cause and effect links are carefully explained. Though not rated within this matrix, the findings section is text heavy and it is recommended, for ease of reading, to include more tables, images and infographics. The conclusions are well-written and provide a thorough understanding and analysis of the context and numerous programmes and initiatives supported by UNFPA, including their cross-cutting elements. Recommendations are organized by their strategic or programmatic components and provide operational considerations for their implementation. The operational implications sufficiently explain the human and technical resources required for their implementation, though the financial implications are not made clear. In addition, all recommendations are oriented towards the country office, and could more clearly reference specific owners/managers in areas of resource mobilization, partnerships, etc. The analysis briefly took into account the extent that the CP specifically considered persons with disabilities, and this was reflected in the conclusions. The recommendations take up PWD within the broader context of LNOB.

**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 assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)*

<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
--	----------------------	-------------------	------------------

*To ensure the report is comprehensive and user-friendly*

<b>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)?</b>	Yes	The report is logically structured and easy to understand.
<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>	Yes	The report falls within the page limits for CPEs.
<b>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)?</b>	Yes	All of the required annexes are included and referenced within the evaluation report.

*Executive summary*

<b>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?</b>	Partial	The executive summary includes all necessary elements with sections on purpose/overview, evaluation approach and users, methodology, findings, conclusions and recommendations. Some of the findings and conclusions presented are vague and general, limiting the summary's use as a stand-alone section. For example, the information presented under effectiveness on the conversion of indicators was general and did not note which indicators lagged or exceeded targets. Gaps are somewhat filled within the conclusions.
<b>5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</b>	Yes	The executive summary is within the 5-page limit.

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
----------------------------------	----------------------	-------------------	------------------

*To ensure that the evaluation is put within its context*

<b>1. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	The complexity of the country context is clearly explained, noting political, security and economic challenges as well as population dynamics.
<b>2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?</b>	Yes	The evaluation report summarizes the theory of change of the country programme in one graphic (Figure 4), and throughout the findings, recommendations for rephrasing and clarifying indicators are made, as relevant.

*To ensure a rigorous design and methodology*

<p><b>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?</b></p>	Yes	The evaluation matrix is complete, including evaluation questions, assumptions, indicators, data sources and methods of data collection. Summarized findings are also presented and organized by evaluation question and data collection method (documents, KIIs, survey).
<p>4. Are the tools for data collection described and their choice justified?</p>	Yes	The evaluation process and methods were clearly explained and included desk review, small group and individual online interviews, observations and an online survey.
<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>	Partial	A comprehensive stakeholder map is included in annex 6, organized by outcome area (e.g. adolescents and youth, gender equality and women's empowerment) and type of stakeholder (e.g. government, non-governmental, etc.). A table on the distribution of respondents reached through interviews is also presented, along with a chart showing the distribution of online respondents by UNFPA programme components. The process for consulting stakeholders across evaluation phases is explained. However, the stakeholder mapping and sampling frame does not clearly specify the extent to which persons with disabilities or their representative organizations were engaged in the country programme or evaluation.
<p>6. Are the methods for analysis clearly described for all types of data?</p>	Yes	The methods of analysis are listed and include document analysis, contribution analysis, content analysis, descriptive statistics and process mapping and visualization.
<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>	Partial	Two methodological limitations were described, most notably the inability to reach end beneficiaries at oblast level considering the COVID-19 pandemic. To mitigate this, it was said that other stakeholders were asked to speak on their behalf, as only 4 beneficiaries were ultimately able to be interviewed. The evaluators note this did not significantly impact the evaluation process. However, the impact on the evaluation process may be under-stated, especially given the proposed methodology included the mainstreaming and application of the 'leave no one behind' policy and integration and consideration of disability-related perspectives, as well as specific evaluation questions enquiring about the needs of the most vulnerable. In that sense, it is recommended that the mitigation methods further elaborate on how the methodology was affected/changed as a result of this limitation. For example, by further clarifying the types of CSOs interviewed and how they may have filled gaps in information related to beneficiaries.
<p>8. Is the sampling strategy described?</p>	Yes	The evaluation applied purposive and convenience sampling techniques for all data collection methods, based on the stakeholder mapping.
<p>9. Does the methodology enable the collection and analysis of disaggregated data?</p>	Yes	The stakeholder lists are disaggregated by sex (male/female) and stakeholder type/level (UNFPA, academia, implementing partners, etc.). The mixed methods methodology and indicators within the evaluation matrix reflect an intent to collect and analyze disaggregated data.
<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>	Partial	As noted, the mixed methods methodology and indicators within the evaluation matrix reflect an intent to collect and analyze disaggregated data, however the questions within the data collection tools, and specifically the survey, do not appear to gather any identifying information (e.g. sex, age, disability status, etc.). The evaluation questions look at the assessed needs and targeting of various groups, the extent to which human-rights based approaches were used, as well as the extent to which partnership frameworks considered or addressed inequalities in women's participation, roles, etc. As noted above, only four end beneficiaries were consulted and it was not fully explained how this limitation was mitigated. Observations did include visits to a shelter for GBV survivors, amongst other medical and support services for survivors of GBV.
<p><b>3. Reliability of Data</b></p>	<p>Yes No Partial</p>	<p>Assessment Level: <b>Very good</b></p>
<p><i>To ensure quality of data and robust data collection processes</i></p>		
<p><b>1. Did the evaluation triangulate data collected as appropriate?</b></p>	Yes	Triangulation is discussed in the methodology section as a means of data validation and is evident in the findings section, with multiple methods (KII, surveys, document review) and sources (e.g. number of interviews or average response on the surveys).
<p><b>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</b></p>	Yes	The evaluation clearly identifies the number of persons targeted for both key informant interviews and surveys, as well as the actual numbers reached in the field disaggregated by the gender of respondents. Overall, 86 KIIs were conducted across government, NGO, donors and UN agencies and 35 survey responses were received. Although the survey response rate is low (28%), the evaluators demonstrated how the respondents covered all stakeholder types and UNFPA programme component areas.
<p>3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</p>	Partial	There is evidence in the discussion of methodology and in the data collection tools that data was collected with sensitivity to issues of discrimination and other ethical considerations. The methodology section notes that cultural and religious sensitivities existing in Ukraine were reviewed and considered, however it does not note which sensitivities were found (if any) and how they were considered in sampling and data collection. It is generally noted that respondents availability for interview was considered, and confidentiality, privacy and anonymity were guaranteed and respected.
<p><b>4. Analysis and Findings</b></p>	<p>Yes No Partial</p>	<p>Assessment Level: <b>Very good</b></p>
<p><i>To ensure sound analysis and credible findings</i></p>		

1. Are the findings substantiated by evidence?	Yes	Findings are well-substantiated with triangulated evidence, with multiple methods (KII, surveys, document review) and sources (e.g. number of interviews or average response on the surveys) referenced for key findings.
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretations are carefully explained.
3. Is the analysis presented against the evaluation questions?	Yes	The analysis is presented against the evaluation questions, with summary findings for each question presented prior to the more in-depth analysis.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Cause and effect links are carefully explained. For example, the report highlights results of the National Programme Pact for Youth 2025, which was funded by UNFPA, including stakeholder perspectives on programme successes, the number of employers joining PACT, and the resulting internships that were provided through this partnership. In addition, campaigns were run by UNFPA to transform perceptions of GBV and data was presented on slight positive shifts in perceptions/non-acceptance of violence. Although there was not an explicit focus on unintended outcomes, evaluators looked closely at the performance of the CP during covid and noted the accomplishments, particularly of communications activities, during this time.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis does present some indicators disaggregated by sex, age and disability status, as available. For example, it was highlighted how two UNFPA mobile clinics provided life-saving medical assistance and referral to specialized services to over 7,900 people (including 3,298 elderly people, 395 people with disabilities and 113 children) from 58 remote settlements.
6. Is the analysis presented against contextual factors?	Yes	Contextual factors influencing and inhibiting results achievement are sufficiently explained, including staff turnover, political instability and the COVID-19 pandemic.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Cross-cutting issues of equity, disability inclusion, gender equality and human rights are adequately considered; for example, assessing the extent to which these issues are mainstreamed into national policies and textbooks, and assessing the extent to which cross-cutting issues have been considered and addressed through programming, research, partnership decisions and advocacy work.
<b>5. Conclusions</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions are directly linked to the findings/evaluation questions.
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 well-written and provide a thorough understanding of the context and numerous programmes and initiatives supported by UNFPA, including their cross-cutting elements.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions are based on well-substantiated evidence presented in the findings, and are therefore considered to be unbiased. They summarize and present a balance of both positive and negative findings.
<b>6. Recommendations</b>		
	Yes No Partial	Assessment Level: <b>Good</b>
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	The recommendations are directly linked to multiple conclusions.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	Recommendations are organized by their strategic or programmatic components and provide operational considerations for their implementation. The operational implications sufficiently explain the human and technical resources required for their implementation, though the financial implications are not made clear. In addition, all recommendations are oriented towards the country office, and could more clearly reference specific owners/managers in areas of resource mobilization, partnerships, etc.
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 sufficiently address areas for filling gaps or building on strengths related to cross-cutting issues.
4. Are the recommendations prioritized?	Yes	The recommendations are prioritized as 'medium' or 'high'.
<b>7. Gender</b>		
	0 1 2 3 <sup>(**)</sup>	Assessment Level: <b>Very good</b>
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) <sup>(*)</sup></i>		

<p>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>	<p><b>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)</b> 1- There is not a specific objective related to the assessment of human rights and gender equality, except in that these considerations are mainstreamed into the country programme results.</p> <p><b>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)</b> 3- There is not a standalone criterion related to gender and human rights, though considerations are mainstreamed into the evaluation questions.</p> <p><b>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)</b> 2- There are multiple evaluation questions, assumptions and indicators which specifically address GEEW.</p> <p><b>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)</b> 2 - The evaluation findings do reflect on the extent to which GEEW results/data are available, though a clear evaluability assessment on HRGE elements is not included as part of the evaluation design process.</p>
<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	<p><b>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)</b> 2 - The methodology specifies how GEEW is addressed in the methodology and data is disaggregated, as available and relevant. The data collection tools, specifically the survey, could have been more explicit in gathering demographic data in order to facilitate additional gender analyses amongst institutional stakeholders.</p> <p><b>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)</b> 2 - The evaluation uses mixed methods, however doesn't clearly explain the mitigation methods for reaching so few beneficiaries, and/or how this may have affected the methodological approach and findings.</p> <p><b>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)</b> 3 Triangulation is clearly applied and diverse data sources are consulted (including observations of interventions/service centers).</p> <p><b>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)</b> 2 - As previously noted, there are some gaps in explanation of how the most vulnerable were engaged/consulted as part of this evaluation. For example, whether specific CSOs were sampled in order to speak on behalf of the most vulnerable.</p> <p><b>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3)</b> 2 Ethical standards</p>
<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>	<p><b>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)</b> 3 - The background section and findings provide a clear intersectional analysis of social groups affected by various issues as well as a mapping of relevant normative instruments and policies.</p> <p><b>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)</b> 3- The findings section clearly disaggregates data by age, sex, disability status, and location, as relevant and available.</p> <p><b>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)</b> 1 - While the unanticipated effects of the intervention are not clearly identified, there are some results presented which are not included as indicators within the country programme's logical framework.</p> <p><b>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)</b> 3 - There are numerous recommendations targeted at addressing GEEW issues, especially considering the most prominent programme component in Ukraine is preventing and responding to gender-based violence. Recommendations consider and propose gender transformative approaches for changing norms, policies and other institutional structures, such as the more systematic engagement of men.</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			
<b>Total scoring points</b>	<b>89</b>	<b>11</b>		

Overall assessment level of evaluation report	<b>Very good</b>
<b>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.</b>	
The evaluation does well to integrate cross-cutting issues, presenting differential results for various groups, as relevant (for example, the elderly) and also presenting disaggregated data on persons targeted/reached through some interventions (e.g. by age, disability status).	
<b>Consideration of significant constraints (e.g. COVID-19 or civil unrest)</b>	
The quality of this evaluation report has been hampered by exceptionally difficult circumstances:	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No
If yes, please explain:	