UNFPA Evaluation Quality Assessment Grid

Version: May 2024

REPORT RATING SUMMARY			
Overall Rating		**68%**	**Satisfactory**
● ● ● ● ●	Excellent	5	
● ● ● ●	Highly Satisfactory	4	
● ● ● -	Satisfactory	**3**	**The report meets UNFPA/UNEG standards for evaluation reports, but some indicators are inadequately addressed or missing. Decision makers may use the evaluation with some confidence.**
● ● - -	Fair	2	
● - - -	Unsatisfactory	1	

REPORT DETAILS	
Title of the evaluation report	UNFPA Angola 8th Country Programme (2020-2022) Country Programme Evaluation
Region	ESA
Country	Angola
Year of report	2022
Business Unit/programme country *(managing evaluation)*	UNFPA Angola Country Office
Date of assessment review (dd/mmm/yyyy)	juin 2024
Name of assessment review firm	IOD PARC
CLASSIFICATION OF EVALUATION REPORT	
Primary SDG(s) covered *(list provided below)*	**3. Good Health and Well-being; 5. Gender Equality**
UNFPA Strategic Plan areas covered *(lists provided below)*	
Three transformative results	1, 2, 3
Six outputs	1. Policy and accountability, 2. Quality of care and services, 3. Gender and social norms
Six accelerators	1. Human rights-based and gender-transformative approaches
Organizational effectiveness and efficiency	No
Humanitarian evaluation	No
Evaluation evaluand *(e.g. country programme/intervention/policy/thematic area)*	Country Programme
Evaluation type *(e.g. formative, summative, developmental)*	Summative and formative
Geographic scope *(e.g. global, regional, national)*	National

EQA Summary: *The rater will provide top line issues for this evaluation relevant for feedback to senior management (strengths and weaknesses), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight good practice/added value elements and the level of complexity of the evaluation. The rater should also highlight how cross-cutting issues were addressed in the report. Considerations of significant constraints (e.g. humanitarian crisis or political turmoil) should also be highlighted here.*

This is an evaluation of UNFPA Angola's 8th Country Programme (2020-2022) Country Programme. The key strengths and weaknesses of the Evaluation Report are as follows:

Strengths and Weaknesses:

• **Section A: Executive Summary –** The Executive Summary is clear and concise. It can serve as a standalone document for decision making. It includes all necessary components with the exception of the scope.

• **Section B: Background –** There is a clear description of the context for the Country Programme including the political context, socio-economic context, education level and demography of Angola. This is well done. Where it can improve is by providing more information on the Country Programme (in the main body or Annex) as it is difficult to understand how it is implemented. It is also difficult to understand who the key stakeholders are and what their role, interest and/or contribution to the Country Programme is.

• **Section C: Evaluation Purpose –** The purpose of the evaluation is clearly defined including why it is needed at this point in time and its intended use. There is a clear and complete description of the objectives of the evaluation as well as the scope (the latter includes geographic scope, thematic scope and temporal scope).

• **Section D: Evaluation Design and Methodology –** The evaluation matrix clearly presents the evaluation criteria with assumptions, indicators, data sources and data collection methods against each evaluation question. This shows a clear line of inquiry from which the analysis will be based on. Where it could improve on is sampling as there is no sampling strategy provided in the report (see recommendations and row 66 below). It could also summarise the types of data sources used in the methodology section and list them in a bibliography/reference page. There is also room for improvement in relation to describing and demonstrating ethical and safeguarding considerations.

• **Section E: Evaluation Findings –** The findings are clearly presented and make explicit use of the results framework. Overall, triangulation is evident, but this can be strengthened by providing reference more consistently as it is not always clear where the sources are from.

• **Section F: Evaluation Conclusion –** The conclusions are clearly formulated. However, they represent more a summary of the findings than offering deeper insight and analysis.

• **Section G: Evaluation Recommendations –** Recommendations are clearly formulated and are derived from findings and/or linked to conclusions. Were they can improve is to include guidance for implementation, timeframe for delivery, prioritisation or identification of responsible actors.

• **Section H: Cross-cutting issues –** Overall, cross-cutting issues are integrated in the core elements of the evaluation. It helps when evaluation questions specifically address cross-cutting issues to ensure that this is well captured. It is also evident that GEEW is integrated in the evaluation scope of analysis, evaluation criteria and questions are designed to ensure GEEW-related data is collected. Where it can improve is by providing a sampling strategy (mentioned above) to understand considerations for allowing for the voices and perspectives of a wide range of stakeholders. There is also limited application of an intersectional lens.

Suggestions for future evaluators: *The rater will identify key suggestions to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, examples will be cited to assist evaluation managers in overseeing future evaluations.*

• **Section B: Background –** Although categories of stakeholders are identified in the methodology section and description of the Country Programme, who they are is not clear. In addition, there is a stakeholder map in Annex 6.7 but main body of the report does not reference it. It would be helpful to include a section in the main body that outlines the key stakeholders and also sign-posts to the Stakeholder Map in the Annex. This would provide the groundwork for understanding the sampling process.
• **Section D: Evaluation Design and Methodology –** It is important to include a sampling strategy in the report in order to understand the rationale behind the selection of informants and to ensure it is robust. While there is a short section titled 'sampling' in the Inception Report, it is quite brief. Furthermore, while it describes the sampling as 'purposive sampling', it does not specify the criteria used.
In terms of ethical considerations, it would be helpful if evaluators added a separate section on ethical considerations in the main body of the report or in the Annex, to describe the evaluation team's approach and demonstrate what steps have been taken. This can include a description of processes in place as well as protocols for how the evaluation was undertaken.
• **Section E: Evaluation Findings –** Evaluators could consider using a rubrics to indicate the strength of evidence that includes triangulation within it.
• **Section F: Evaluation Recommendations –** Where possible, recommendations could be accompanied by 'supporting actions' so it is clear what steps can be taken in order to deliver recommendations are that potentially more strategic or transformative. It would also be good to include timeframe for delivery, prioritisation, identification of responsible actors and so on.

SECTION RATINGS

SECTION A:	**EXECUTIVE SUMMARY (weight 5%)**	83%	**Comments on Rating**
Question 1.	**Can the executive summary inform decision-making?**		
i	Is a clear, standalone document useful for informing decision making, (a minimum of 5 pages, up to a maximum of 7 pages). *Note: YES - the executive summary is within the indicated maximum page limit. PARTIAL - the executive summary exceeds the maximum page limit by 1 to 2 pages. NO - the executive summary exceeds the maximum page limit by more than 2 pages.*	Yes	The Executive Summary is succinct yet informative. It can serve as a standalone document. It provides useful information for decision-making. It is four and a half pages in length. Please note that the ToR states that the Executive Summary should be '*a maximum length of 5 pages*' (p.69).
ii	Includes all necessary components of the evaluation report, including: (1) overview of the context and intervention, (2) evaluation purpose, objectives and intended users, 3) scope and evaluation methodology, (4) summary of most significant findings, (5) main conclusions and (6) key recommendations	Partially	It includes the main components of an evaluation report, starting with a background section that offers an overview of the context of the country programme. It includes a section outlining the purpose, objectives and intended audience. The methodology section provides a brief, but sufficient, overview of the approach and methods employed by the evaluation team. Additionally, it presents the main findings, conclusions and recommendations. What is missing is the scope of the evaluation. This would be helpful to incorporate as the Executive Summary is less than 5 pages in length. Furthermore, the findings and conclusions are combined into a single section.
iii	Includes all significant information in a concise yet clear manner to understand the theme, intervention, programme, project and the evaluation.	Yes	The Executive Summary includes key information in a concise and clear manner. It could further elaborate on the country context and the country programme, particularly if the page limit is a maximum of 7 pages. However, as the ToR does specifies a lower page limit, the information provided is sufficient.
SECTION B:	**BACKGROUND (weight 5%)**	60%	**Comments on Rating**
Question 2.	**Is the evaluand (i.e. intervention/policy/thematic area etc. that is to be evaluated) and context of the evaluation clearly described?**		
i	Clear description of the evaluand (e.g. intervention), including: geographic coverage, implementation period, main partners, cost/budget, and implementation status.	Partially	There is a section on the Country Programme, which provides an overview of the key stakeholders involved in its development. It also explains how the Country Programme aligns with relevant national and UN strategies and policies. Additionally, it compares the 7th and 8th Country Programmes, highlighting key changes. There is a section on the financial structure. However, the section could be improved by including a description of what the 8th Country Programme aims to achieve and how it is implemented. Currently, it is difficult to understand the workings of the Country Programme without referring to the Results Framework in the Annex, which this section does not reference. It would be helpful if the section outlined the over-arching goal, followed by a brief description of the three areas of programming, and then sign-post to the Results Framework for more detail. **In addition, while para 43 provides a table that outlines the outcome areas and outputs of, the ToR provides a good description of how the outputs are delivered (see p.4-6).** If there is not enough room in the report, it can be signposted to the Annex.
ii	Clear description of the context of the evaluand (e.g. economic, social and political context, relevant aspects of UNFPA's institutional, normative and strategic framework, cross cutting issues such as gender equality and human rights, disability and LNOB dimensions) and how the context relates to the evaluand (e.g. key drivers and challenges that affect the implementation of the intervention/policy/thematic area	Yes	There is a clear description of the context for the Country Programme. This includes socio-political context, economy, education and demography of Angola. It provides a section on reproductive health, which is central to the country programme. It describes gender equality and women's empowerment for Angola including how it ranks in the World Gender Inequality Index and Gender Parity Index. It also outlines issues including gender-based violence (GBV). The context is helpful for understanding UNFPA's strategic response, which it also describes.

iii	Linkages drawn between the evaluand and the ICPD benchmarks and SDGs relevant targets and indicators.	Yes	Linkages are drawn to the relevant SDG and, to a lesser extent, ICPD benchmarks. On the latter, more information is included in the Findings section (part of EQ1) than the background and context sections, where it is mentioned as part of the purpose for the evaluation (p.2).
Question 3.	**Are key stakeholders clearly identified and analysed?**		
i	Clear identification of key stakeholders which should include implementing partner(s), development partners, rights holders, and duty bearers among others; and of linkages between them (e.g., stakeholder map).	Partially	Key stakeholders are not clearly identified in the report. Although categories of stakeholders are identified in the methodology section and description of the Country Programme, the specific key stakeholders remain unclear. There is a stakeholder map in the Annex (6.7) that identifies key partners against each of the outcome areas. This table as provided in the ToRs and does not appear to have been updated. Furthermore, the main body of the report does not reference the stakeholder map. It would be helpful to include a section in the main body that outlines the key stakeholders and also sign-posts to the Stakeholder Map in the Annex. This would provide the groundwork for understanding the sampling process. Additionally, it would also be good to identify who the key rights holders are. Although there is a column in the table in Annex 6.7 labelled 'rights holders', it remains unpopulated.
ii	Stakeholders are analysed to understand their specific rights, duties, needs, interests, concerns, and potential impact on the evaluand.	No	This is missing from the report. It would be helpful to understand not just who the key stakeholders are but how they contribute to the Country Programme and/or their needs, interests and concerns (for both duty bearers and rights holders).
SECTION C:	**EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)**	100%	**Comments on Rating**
Question 4.	**Is the purpose of the evaluation clearly described?**		
i	Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Yes	The purpose of the evaluation is clearly defined as being two-fold: (1) independent assessment for the intended audience and (2) to help inform the design of the next programme cycle (p.6). As such, it meets the criteria for why it is needed at this point in time and its intended use. It may be worth noting that the purpose of the evaluation is described as the first set of objectives and the second set described as 'specific objectives' (please see p.9 of ToRs). As mentioned, the key intended users are incorporated into the first part of the purpose (i.e. *"i. To provide the UNFPA Angola CO, national stakeholders and rights-holders, the UNFPA East and Southern Africa Regional Office [ESARO], UNFPA Headquarters as well as a wider audience..."*). It would be preferable if there was a separate sub-section that outlines who the intended audiences are rather than as part of the purpose itself. It would also be good to differentiate between including primary and secondary users. Ideally, there would be one purpose per evaluation and this one could be *"To provide an independent assessment of the UNFPA Angola 8th Country Programme (2020-2022) in order to broaden the evidence base to inform the design of the next programme cycle."* However, this suggestion may go beyond what is asked in the criteria and, therefore, does not affect the rating.
Question 5.	**Are the objectives and scope of the evaluation clear and realistic?**		
i	Clear and complete description of the objectives of the evaluation, including reference to any changes made to the objectives included in the ToR (if applicable).	Yes	There is a clear and complete description of the objectives of the evaluation (p.2). There are no changes made to the specific objectives included in the ToRs.
ii	Clear and relevant description of the scope (e.g. thematic, geographic, and temporal) of the evaluation, covering what will and will not be covered, as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political, humanitarian or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	There is a clear and relevant description of the scope of the evaluation. This includes the geographic scope (all regions and provinces where UNFPA implemented interventions), thematic scope (all three thematic areas of the 8th CP, COVID-19 response and cross-cutting issues), and temporal scope (period of the CP 2020-2022).
SECTION D:	**EVALUATION DESIGN AND METHODOLOGY (weight 20%)**	58%	**Comments on Rating**
Question 6.	**Are the selected evaluation questions and evaluation criteria appropriate for the purpose of the evaluation and is there clear justification for their use?** *Note: UNFPA evaluation standards refer to the OECD/DAC criteria such as: relevance, coherence, effectiveness, efficiency and sustainability (not necessarily applicable to all evaluations) and, for country programmes that include circumscribed and limited humanitarian and/or emergency interventions, the criteria of coverage and connectedness.*		

i	Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluation questions in the final evaluation report are the same as those in the ToRs so no changes were made during inception. This is a missed opportunity as there are ten evaluation questions, which is a high number. Moreover, the ToR emphasises that the questions are 'indicative and preliminary' and that evaluators are expected to develop the final questions with key stakeholders (see p.11 of ToRs). In similar future evaluations, evaluators can be further encouraged during inception (as the ToRs already encourage this) to explore whether the questions could be consolidated and if some could serve as sub-components. For example, EQ2 and EQ9, which are both related to Covid-19. Overall, the evaluation questions and evaluation criteria are appropriate and aligned to the objectives of the evaluation.
ii	Evaluation matrix clearly presents the evaluation criteria used as well as the corresponding evaluation questions, indicators, lines of inquiry, benchmarks, assumptions, source of data, methods for data collection and analysis, and/or other processes from which the analysis can be based, and conclusions drawn.	Yes	The evaluation matrix (Annex 6.1) uses the template provided in the ToRs. It clearly presents the evaluation criteria as well as corresponding evaluation questions. Against each question are assumptions to be tests and these include indicators, data sources and method for data collection. This shows a clear line of inquiry from which the analysis is based on and conclusions drawn.
Question 7.	**Is the theory of change, results chain, logical framework, or equivalent framework well-articulated?**		
i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Partially	The Country Programme's results framework is provided in Annex 6.3. It includes the overarching goal, the intended outcomes for the thematic areas (including target and baseline), outputs (including indicators and targets) and supporting activities. **However, it does not include the nested ToCs for each outcome, which are provided in the ToR and this evaluation is expected to assess.**
ii	Causal relationships between the various elements (e.g. outcomes, including the three or relevant Transformative Results, outputs) of the theory of change, results chain or logical framework are presented in narrative and/or graphic form).	Yes	The causal relationships are presented in graphic form in Annex 6.3 as per the Country Programme Results Framework (described above in row 60). There is also a brief narrative description in the context/background section of the Country Programme (p.18-19). It would be helpful if there was clearer signposting to the Annex from the narrative although it is possible to make the links through the content page.
iii	Comprehensive analysis and assessment of the theory of change, results chain or logical framework, and if requested in the ToR, it is retrofitted/reconstructed by the evaluators.	Partially	The methodology provides an explanation for how contribution analysis was used for the evaluation, indicating an analysis and assessment of the theory of change (p. 4). It is also evident from the evaluation matrix that assumptions for each of the questions have been identified for assessment. Where improvement could be made is for evaluators to include the 'nested theory of change' or descriptions of what these are. The methodology mentions that the nested theories were developed for each outcome that include assumptions to the causal links. However, these are not included and it is unclear how the assumptions in the evaluation matrix were identified and what was identified from the nested theory of change for each outcome area.
Question 8.	**Does the report specify adequate methods for data collection, analysis, and sampling?**		
i	Evaluation design and set of methods are clearly described, and are relevant and robust for the evaluation's purpose, objectives and scope, including the use of AI in the evaluation process if applicable.	Partially	The evaluation design and methodology is provided but it is fairly brief (largely condensed into one paragraph - see para 8 on p.4). This shows how contribution analysis was used, which is appropriate for meeting the evaluation's purpose, objectives and scope. **In terms of data collection methods, it clearly describes interviews and FGDs that took place online due to COVID restrictions. The stakeholders interviewed are also described (para 9) and the tools are provided in the annex. It would be helpful if this section sign-posted to the interview tools in the Annex and also clarified if the same tools were used for the FGDs. The key gap to assess the extent that the evaluation design is robust is the sampling strategy, but this is covered in more detail in row 66.** Lastly, the evaluation report does not mention the use of AI in the evaluation process.

ii	Data sources are all clearly described and are relevant and robust; these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR).	Partially	**Para 9 includes a list of the types of data sources used. This is also reflected in the evaluation matrix, which** includes data sources for each evaluation question. This includes a wide range of documentary evidence, key informant interviews (Annex 6.2 includes list of people interviewed) and FGDs. These are relevant for the evaluation and the types of sources are robust (though sampling strategy is further elaborated in row 65). The original evaluation design proposed the use of a survey. However, this did not go ahead and a justification is provided under limitations (p.6). As the survey was not used, the list of surveys (Annex 6.6) can be removed from the Annex. In addition, evaluators can consider including a bibliography/reference page for all secondary sources used. A significant number of documents cited in the evaluation matrix are abbreviated and it would be helpful to have a reference page that lists all the secondary sources used. The Findings section includes a number of secondary sources used in the footnotes and a bibliography will help bring them together.
iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	No	There is no sampling strategy provided in the report. Table 2 (p.6) provides a list of categories of stakeholders interviews alongside the numbers used and the data collection method. However, there is no explanation for why and how these stakeholders were identified (i.e. criteria for selection). It is also unclear what regions these stakeholders covered and why. There is also no description or introduction of the table in the narrative of the report. Without an understanding of how these were arrived at, KIIs and FGDs with 40 individuals seems low for a country programme covering a wide geographic area with a budget of $8.7m. It may be sufficient but it is difficult to assess without any documentation on the rationale. It may also be worth noting that there is a short section titled 'sampling' in the Inception Report. However, it is quite brief. While it describes the sampling as 'purposive sampling', it does not specify the criteria used.
iv	Methods allow for rigorous testing of the theory of change, results chain or logical framework (e.g methods help to understand the causal connections, if any, between outputs and expected outcomes (3TRs).	Yes	Contribution analysis, which this evaluation applies, allows for rigorous testing of the theory of change including the causal connections between outputs an expected outcomes.
v	Clear and complete description of the methods of analysis, including explanability and full disclosure of the use of AI in the evaluation process, if applicable.	Partially	There is a description of the method of analysis in the main body of the report, albeit brief. **However, the evaluation matrix provides clear assessment criteria for how judgements would be made against each evaluation question and corresponding assumption areas (Annex 6.1). In terms of its use of contribution analysis, it could provide a bit more information, drawing from content in the Inception Report. It would also be helpful if the 'nested theories' are described and/or added to the Annex. In addition, it could describe how the key assumptions in the evaluation matrix were arrived at.** The evaluation report does not mention the use of AI in the evaluation analysis.
vi	Clear and complete description of limitations and constraints faced by the evaluation in its data collection and analysis, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	There are two items included in the limitations section of the report. One was related to limitations relating to Covid-19 and the team's inability to travel and the second to the online survey not being used. It could further elaborate on the limitations mentioned elsewhere in the report. For instance, data quality or data availability issues is evident in Annex 6.4, which states that results achieved against target outcome indicators were not available.
Question 9.	**Are ethical issues and considerations described?** **The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:**		
i	Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or UNEG Ethical Principles.	Partially	The evaluation report makes explicit reference to the UNEG norms and standards and UNEG Ethical guidelines for Evaluation, as well as UNFPA evaluation guidelines. However, it does not specify how it adhered to ethical principles and standards.
ii	Clear description of ethical issues and considerations (e.g. respect for dignity and diversity, fair representation, confidentiality, and avoidance of harm) that may arise in the evaluation, safeguard mechanisms for respondents (e.g. parental consent forms for adolescents, compliance with codes for vulnerable groups; WHO standards of safe data collection on GBV) and ethical considerations in the use of AI as applicable (e.g, transparancy of use, explainability, privacy, data protection, accuracy, human rights). If AI is used in the evaluation, there should be transparency and disclosure on the ethical and responsible use of AI in the report.	Partially	The evaluation report references UNEG ethical guidelines for evaluations. However, there is limited description of ethical issues and considerations in the evaluation report. It would be helpful if evaluators added a separate section on ethical considerations in the main body of the report or in the Annex to describe the evaluation team's approach and demonstrate what steps have been taken. This can include a description of processes in place as well as protocols for how the evaluation was undertaken. For instance, the Annex includes the questions asked in interview (Annex 6.5) but it does not include the protocols e.g. what is communicated to informants on issues of confidentiality and so on. It could also include protocols to ensure safeguarding mechanisms for respondents. As FGDs also involve girl beneficiaries, IRB approval may have required but it unclear if this was sought and, if not, what the justification is.
Question 10.	**Does the evaluation incorporate innovative practice that adds value to the evaluation process?**		

i	Innovation practice is used to improve the quality of evaluation process. This could include efforts to optimize the evaluation process (e.g., use of AI or new technology for data gathering, content analysis, outcome harvesting among others), or components introduced to enhance inclusion and participation in the evaluation processes (e.g. a youth steering committee), or ways of sharing of evaluation results.	Not Rated	While nested theories of change are not novel, its application can be considered innovative in its use to enhance the quality of evaluation. Unfortunately, not enough detail is provided for how it was applied. For this reason, this section is rated as '**Not Rated**'.
SECTION E:	**EVALUATION FINDINGS (weight 25%)**	67%	**Comments on Rating**
Question 11.	**Do the findings clearly and adequately address all evaluation questions and sub-questions?**		
i	Findings are presented clearly and provide sufficient levels of evidence to systematically address all the evaluation's questions	Yes	Findings are presented clearly. They are structured according to OECD and ALNAP criteria and evaluation questions. Overall, the findings section provides sufficient levels of evidence to systematically address all the evaluation questions.
ii	Explicit use of the evaluand's theory of change, results chain, logical framework in the formulation of the findings.	Yes	The findings make explicit use of the Country Programme's results framework in the formulation of the findings. This is particularly evident for questions related to effectiveness where findings were presented against each outcome area.
Question 12.	**Are evaluation findings derived from credible data sources as well as a rigorous data analysis?**		
i	Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident using multiple data sources.	Partially	The evaluation uses credible forms of qualitative and quantitative data including a range of secondary sources. It also presents output and outcome-level data. Overall, triangulation is evident in the findings section such as use of secondary sources with primary data (such as key informant interviews and focus group discussions). To strengthen this component, evaluators could be encouraged to provide references more consistently as it is not always clear where their sources are from, with some sections stronger than others. For example, para 172 (p.53) indicates the different sources yet para 203 (p.59) includes a list of trainings with number of recipients but it does not state what the source is and if/how it has been triangulated. What evaluators can consider is using a rubrics to indicate the strength of evidence that includes triangulation within it. In addition, the lack of sampling strategy and clarity about the sampling size remains a concern about the credibility of the data.
ii	Findings are clearly supported by the evidence presented, both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Yes	Findings are clearly supported by evidence that includes both positive and negative. There are high-level 'preliminary findings' statements - 17 in total. These findings statements are well crafted as they are succinct yet nuanced and, often, include positive and negative findings. Each of the findings statements are supported by evidence. There is a very clear evaluation matrix with assessment criteria for sub-components/assumptions tested against each evaluation question (Annex 6.1).
iii	Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from outputs to high level results).	Partially	Causal factors leading to achievement and non-achievement of results are identified. As this is a theory-based evaluation, the findings analyse the contributing factors and to, to some extent, the logical chain (i.e. progression or not from output to outcome). Where it can improve is to discuss or elaborate on the achievements or non-achievements at output level (provided in the tables) but this is not always discussed systematically or at all. It is unclear how the achievements or non-achievements at output levels affect the higher level results (Outcome 2 (under effectiveness EQ3) is one example).
Question 13.	**Does the evaluation assess and use the intervention's Results Based Management elements?**		
i	Assessment of the adequacy of the intervention's planning, monitoring, and reporting system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	No	Conclusion 11 describes the challenges in building an integrated M&E system. However, it is unclear which findings section this conclusion is drawn from and what evidence support it. **As this question relates to the findings section, the report does not satisfy this criteria.**
SECTION F:	**EVALUATION CONCLUSIONS (weight 10%)**	75%	Comments on Rating
Question 14.	**Do the conclusions clearly present an unbiased overall assessment of the evaluand?**		
i	Conclusions are clearly formulated and present unbiased summative statements that respond to the evaluation questions.	Yes	The conclusions are clearly formulated - they are numbered and there are 12 in total. They present summative statements that respond to the evaluation questions. In fact, nearly each conclusion aligns with an evaluation question.
ii	Conclusions are well substantiated and derived from findings and add deeper insight and analysis beyond the findings.	Partially	**The conclusions are largely derived from the findings, this is evident from the number of conclusion paragraphs correspond to the evaluation questions in the findings section (e.g. Conclusion 1 relates to EQ1, Conclusion 2 relates to EQ2, Conclusion 7 relates to EQ4 etc). This is with the exception of conclusion 11. As a result, the conclusions** represent more a summary of the findings than offering deeper insight and analysis. The conclusion is an opportunity to take a step back and to look at the broader picture.

Question 15.	**Are lessons learned identified? [N/A if lessons are not referenced or requested in ToR]**		
i	Lessons learned are derived from the findings and are well substantiated with practical, illustrative examples.	Not Rated	There is no separate section on lessons learned. They are not explicitly requested in the ToRs.
ii	Lessons learned are clearly presented and provide actionable insights on the positive aspects of the evaluand as well as any areas of improvement.	Not Rated	As above, there is no separate section on lessons learned as they are not explicitly requested in the ToRs.
SECTION G:	**EVALUATION RECOMMENDATIONS (weight 15%)**	63%	**Comments on Rating**
Question 16.	**Are recommendations well-grounded and articulated?**		
i	Recommendations are clearly formulated and logically derived from the findings and/or conclusions.	Yes	The recommendations are clearly formulated and are derived from the findings and/or are linked to the conclusion. Each recommendation includes sign-posting to the conclusion that it is related to.
ii	Recommendations are useful and actionable for primary intended users. Specific guidance is provided for its implementation (e.g. actions, deadlines, responsible actors), as appropriate.	Partially	Recommendations identify areas for improvement and what can be done. Where possible, recommendations could be accompanied by 'supporting actions' so it is clear what steps can be taken in order to deliver recommendations are that potentially more strategic or transformative. There is no specific guidance provided for tha implementation. This would be good to include timeframe for delivery, prioritisation, identification of responsible actors and so on.
iii	Process for developing the recommendations is described, and includes the Involvement of key stakeholders (e.g. evaluation reference group members), including those who will be affected by the recommendations.	Partially	The process for developing of recommendations is not described. However, it states that: *"these are preliminary recommendations that will be adjusted at the end of the year after the assessment of the results achieved and a workshop with the main stakeholders"* (p.62).
iv	Recommendations are clearly articulated and prioritized based on their importance, urgency, and potential impact.	Partially	Recommendations are clear but they are not prioritised based on their importance, urgency and potential impact.
SECTION H:	**REPORT STRUCTURE AND PRESENTATION (weight 5%)**	75%	**Comments on Rating**
Question 17.	**Does the evaluation report include all required information?**		
i	Opening pages include: Name of evaluation and/title of evaluation, timeframe of the evaluation, date of report, location of evaluand, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents (including, as relevant, tables, graphs, figures, annexes)-; list of acronyms/abbreviations.	Yes	The opening pages include all the required information. The evaluation report includes the name of the evaluation (as indicated on row 13). and the timeframe of the evaluation and data of the report on the cover page. The location of the Country Programme is included in the title alongside a map on the second page. It also includes the names of the evaluators and their roles on the assignment and a logo of UNFPA to indicate the organisation commissioning the evaluation. A comprehensive table of contents is included including tables, graphs, and images. There is also a page of abbreviations.
ii	Annexes include, if not in body of report: terms of reference, evaluation matrix, list of respondents, results chain/ToC/logical framework, list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology (e.g. inception report), case study reports.	Partially	The Annexes include key information but it is not complete. It includes the evaluation matrix, list of respondents, results framework, data collection instruments (although it includes survey instrument, which was not rolled out in the evaluation). The terms of reference are not included in the Annex but as they are 96 pages in length, a separate Annex may be appropriate. **The Annex is also missing the ToC including nested ToCs (provided in the ToR)** and a list of documentary evidence/bibliography. Evaluators can also be encouraged to use the Annexes to include protocols used in the evaluation, in addition to the survey instruments, to illustrate how confidentiality and consent was communicated to respondents. There are also other areas that could be further elaborated and documented, such as ethical considerations and methodology, as suggested in some comments above.
Question 18.	**Is the report logically structured and of reasonable length?**		
i	The report has a logical structure that is easy to identify and navigate (for instance, with numbered sections, clear titles, well formatted).	Yes	The report has a logical structure that is easy to identify and navigate. The content page is helpful for navigating. It also has numbered paragraphs in addition to numbered sections and sub-sections. It has clear titles and is well formatted.
ii	Structure and length accords to UNFPA guidelines for evaluation reports; it does not exceed number of pages that may be specified in ToR. *Note: Maximum pages for the main report, excluding executive summary and annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations and 50 for other types of evaluations)*	Yes	The length of the report is 64 pages (excluding the executive summary), which is within UNFPA guidelines for Country Programme evaluations.
Question 19.	**Is the report well presented?**		
i	Report is easy to understand (written in an accessible way for the intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	The report is fairly easy to understand. However, it would benefit from additional proof-reading as there some spelling mistakes (e.g. 'creatn' on para 105 and 'e' on para 209) and some issues with spacing within sentences. The page numbers also need addressing as the executive summary jumps from p.7 to p.2. The main section of the evaluation starts with page 2 rather than page 1. These are fairly minor but there are few of them in the report that could be tightened with proof-reading.

ii	Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Partially	There are some visual aids such as maps and tables. These are clearly presented and labelled. However, they are not always referenced in the text (as mentioned in comments above). For example, Table 2 on category of stakeholders is not referenced or discussed in the text. This is also the case in other sections.
SECTION I:	**CROSS-CUTTING ISSUES (weight 10%)**	67%	**Comments on Rating**
Question 20.	**Are cross cutting issues - in particular, human rights-based approach, gender equality, disability inclusion, LNOB - integrated in the core elements of the evaluation (e.g. evaluation design, methodology, findings, conclusions and recommendations)?**		
i	Evaluation's data collection methods designed to capture the voices/perspectives of a wide range of stakeholders including right holders, marginalized and vulnerable persons, young people, people with disabilities, migrants or refugee populations, indigenous communities, and other persons that are often left behind.	Partially	The evaluation's data collection methods allow for the voices/perspectives of a wide range of stakeholders to be included through the use of KIIs and FGDs. However, as the sampling strategy was not included, the rationale for selecting two FGDs involving 15 girl beneficiaries in total (p.5) is unclear (please see row 66 on sampling for more). It also does not explain where data is collection and triangulated from secondary sources alone in relation to capturing the voices of a wide range of stakeholders.
ii	Evaluation questions address cross cutting issues, such as human rights-based approach, gender equality, disability inclusion, LNOB, social and environmental standards as appropriate.	Yes	The evaluation questions address cross cutting issues. There is a specific question on human rights, gender mainstreaming and LNOB. In addition, it is evident from the evaluation matrix and the findings section that areas such as gender equality and disability inclusion are addressed in other questions as well.
iii	Data is disaggregated by population groups (e.g. persons with disability, age, gender, etc.) where there are implications related to UNFPA's portfolio/interventions for these population groups; differential results are assessed (distribution of results across different groups).	Yes	The evaluation report does not disaggregate primary sources such as by age, gender or disability. However, it does specify that the FGDs were conducted with girl beneficiaries. In terms of its findings, it mentions UNFPA's disaggregation of data (e.g. para 192 on p.57). UNFPA is also working on providing capacity building of public institutions at various levels to generate and use of disaggregated data. This is an part of output 1 under Outcome 4 on Population Dynamics. It is also explored under EQ3 and EQ8 by the evaluation team. In addition, preliminary finding 11 states that UNFPA has played a key role in the implementation of pre-census activities (p.44-45).
iv	Intersectional lens is applied in the data analysis, looking at various and multiple forms of exclusion and discrimination (and how they overlap with each other) and how this may impact the performance or results of the evaluand.	Partially	The application of an intersectional lens is limited as challenges are discussed separately (e.g. by disability, gender etc.). However, the findings related to EQ4 of e.g. girl who are trafficked, are at risk of HIV or GBV (p.48) to raise the need for more targeted approach to reach those most vulnerable.
v	Findings, conclusions and recommendations, address cross-cutting issues such as equality and vulnerability, disability inclusion, leave no-one behind, social and environmental as relevant.	Yes	The findings, conclusions and recommendations address cross-cutting issues such as equality, disability inclusion and LNOB.
vi	Inclusion of young people in the evaluation team and/or Reference Group [N/A if not requested in ToR]	No	The ToR requests that a young and emerging evaluator is included in the evaluation team. However, it is unclear from the final report if at least one of the three named evaluators are young evaluators. One carries the title 'Youth and Gender Specialist' but it is not clear if this is a young evaluator. Evaluators are encouraged to include a brief description of the team composition in the main body or Annex in order to be able to answer this criteria. Alternatively, evaluators can include "Young and emerging evaluator" in the title as described in the ToR.
Question 21.	**Does the evaluation meet UN SWAP evaluation performance indicators?** *Note: this question will be rated according to UN SWAP standards with detail provided below*	6	**Comments on Rating**
i	GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	GEEW is integrated in the evaluation scope of analysis, evaluation criteria and questions are designed in a way to ensure GEEW-related data will be collected. There is a stand alone question on GEEW but gender considerations are also evident in the findings section and the evaluation criteria.
ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	The methodology refers to Integrating Human Rights and Gender Equality in Evaluation document and evaluators describe that the evaluation examined the level of gender mainstreaming across the country programme, in line with UN's gender mainstreaming methodology. There were also FGDs with girl beneficiaries and the evaluation matrix includes gender considerations. However, not enough detail is provided in the methodology to demonstrate that gender considerations are made to ensure data collected is disaggregated by gender, a sampling strategy (elaborated on above), and the protocols used to demonstrate ethical standards considered throughout the evaluation.

iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	Overall, the evaluation findings, conclusion and recommendations reflect a gender analysis. While the evaluation does not include a background section that includes intersectional analysis, it does reference normative instruments related to human rights and gender equality such as Integrating Human Rights and Gender Equality in Evaluation. The evaluation report also provides specific recommendations that address GEEW issues.

SWAP Rating Guidance

i GEEW is integrated in the evaluation scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.
a. 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?
b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?

ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.
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?
b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?
c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?
d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?

iii The evaluation findings, conclusions and recommendations reflect a gender analysis.
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?
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?
c. Are unanticipated effects of the intervention on human rights and gender equality described?
d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?

List of SDGs

1. No Poverty
2. Zero Hunger
3. Good Health and Well-being
4. Quality Education
5. Gender Equality
6. Clean Water and Sanitation
7. Affordable and Clean Energy
8. Decent Work and Economic Growth
9. Industry, Innovation and Infrastructure
10. Reduced Inequality
11. Sustainable Cities and Communities
12. Responsible Consumption and Production
13. Climate Action
14. Life Below Water
15. Life on Land
16. Peace, Justice and Strong Institutions
17. Partnerships for the Goals

Three transformative results
1. Ending unmet need for family planning
2. Ending preventable maternal deaths
3. Ending gender-based violence and harmful practices
Six outputs
1. Policy and accountability
2. Quality of care and services
3. Gender and social norms
4. Population change and data
5. Humanitarian action
6. Adolescents and youth
Six accelerators
1. Human rights-based and gender-transformative approaches
2. Innovation and digitalization
3. Partnerships, South-South and triangular cooperation, and financing
4. Data and evidence
5. Leaving no one behind and reaching the furthest behind first
6 .Resilience and adaptation, and complementarity among development, humanitarian and peace-responsive efforts