

UNFPA Evaluation Quality Assessment Grid

Version: May 2024

REPORT RATING SUMMARY			
Overall Rating		67%	Satisfactory
● ● ● ● ●	Excellent	5	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.
● ● ● ● ●	Highly Satisfactory	4	
● ● ● ● ●	Satisfactory	3	
● ● ● ● ●	Fair	2	
● ● ● ● ●	Unsatisfactory	1	

REPORT DETAILS	
Title of the evaluation report	
Country Programme Evaluation (CPE) of the 9th Government of Uganda/UNFPA Country Programme (2021– 2025) in Uganda	
Region	
ESA	
Country	
Uganda	
Year of report	
2025	
Business Unit/programme country (managing evaluation)	
Uganda Country Office	
Date of assessment review (dd/mm/yyyy)	
March 21, 2025	
Name of assessment review firm	
CLASSIFICATION OF EVALUATION REPORT	
Primary SDG(s) covered (list provided below)	
3, 5	
UNFPA Strategic Plan areas covered (lists provided below)	
Three transformative results	
(a) an end to preventable maternal deaths; (b) an end to the unmet need for family planning; and (c) an end to gender-based violence and all harmful practices, including female genital mutilation and child, early and forced marriage	
Six outputs	
(a) policy and accountability, (b) quality of care and services, (c) gender and norms, (d) data and evidence, including on population changes, (e) humanitarian action, and (f) adolescent and youth.	
Six accelerators	
(a) Human rights-based and gender transformative approaches (b) Innovation and digitalization (c) Partnership, South-South and triangular cooperation, and financing; (d) Data and evidence; (e) “Leaving no one behind” and “reaching the furthest left behind first”; (f) Resilience and adaptation,	
Organizational effectiveness and efficiency	
Yes	
Humanitarian evaluation	
Yes	
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.	
Strengths	
• There is a clear and appropriate description of the purpose, objectives and scope of the evaluation. • Evaluation questions are clear and organised against relevant OECD DAC criteria, including humanitarian-specific criteria of coverage and connectedness. • The methodology is robust, with appropriate methods such as KIIs, FGDs and document review, but also includes more innovative approaches, such as the use of outcome harvesting, and most significant change. Additionally, an assessment of integrated healthcare is provided by assessing 81 health facilities and rating their performance. • Findings address each evaluation question systematically and use a range of evidence and data sources providing examples of positive and negative experiences. An example of good practice is the assessment of each ToC and their usage in the formulation of the findings. • Conclusions are clear and linked to findings and recommendations and lessons learned are provided as part of the findings sections. • Recommendations are appropriate and based upon findings and conclusions. • It is clear that this report sought to integrate gender and inclusion from the outset, with consideration of LNOB a number of evaluation questions explicitly relating to those issues.	
Weaknesses	
• The Executive Summary could have been improved for clarity. For example, conclusions could have been presented more clearly, and each recommendations listed separately for clarity. • The context section would have improved if it had included a description of LNOB principles, and cross-cutting issues such as disability. • While it is clear that the evaluators reformulated the ToCs the graphical images of them could be presented more clearly and numbered appropriately in the annexes. • While a stakeholder map is provided at Annex 5, this could provide more detail on the roles and responsibilities of all stakeholders, aligned with CP outputs. Additionally, a fuller analysis of the specific rights, duties and responsibilities of different stakeholders would have improved understanding. • The Evaluation matrix would have benefited from a deeper analysis of assumptions, and how they were developed/ tested, and if it had been presented in accordance with the UNFPA Evaluation Handbook 2024. • The inclusion of a criteria on “coordination” should be further explained. As per the UNFPA Evaluation Handbook issues of coordination should be covered in the OECD DAC criterion of coherence. UNFPA no longer uses a standalone criterion on coordination. • With regard to the methodology, while limitations are provided, they could have been more detailed/ thorough. A missed area which would have significantly improved the methodology, and subsequent sections, would have been explicit disaggregation of primary data, for KIIs and FGDs, and to assess how well outcome harvesting and MSC were applied in practice in the conduct of the evaluation. • Ethical issues and consent, are outlined, but how ethics will actually be applied to this evaluation should be provided, along with specific consideration of safeguarding, and wider ethical considerations. • With regard to findings areas for improvement relate to include explicit analysis of results across groups, and providing an assessment of the CP’s monitoring systems and processes. • With regard to recommendations their utility could be improved if greater clarity was provided on the time frames (in months), if specific detail of responsible parties beyond the CO or “partners” were provided, and if more detail on how recommendations were developed was provided. • The report would benefit from a thorough formatting and proof reading, particularly the annexes. • With regard to integration of gender, and inclusion more widely, this could have been improved if gender has been more clearly integrated across the evaluation. However, full integration of GEWE considerations was hindered by a lack of disaggregated data, and intersectional analysis.	
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.	
Suggestions for future evaluations:	
• Clarify and consistently apply evaluation criteria to avoid discrepancies between narrative, tables, and figures. • Strengthen causal analysis and assumptions by providing deeper, more comprehensive exploration of why results occurred or not. • Improve disaggregated data collection and use, ensuring gender, age, disability, and other relevant variables are systematically analysed. • Apply an intersectional lens to assess overlapping forms of exclusion and their effect on results. • Expand and align stakeholder mapping to show linkages and responsibilities clearly, ensuring completeness and accuracy. • Assess and document M&E systems and tools to evaluate their adequacy for decision-making. • Refine recommendations with specific responsible actors, clear timelines, and actionable steps. • Enhance annexes’ structure and accessibility, ensuring proper labelling, proofing, and removal of inaccessible files. • Increase clarity and usability of visuals by improving readability, labelling, and direct linkage to the narrative. • Integrate gender equality more fully across scope, methodology, data analysis, and findings using a structured gender-responsive approach.	
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).	Yes	The executive summary is a clear and concise standalone document, useful for informing decision-making and is 5 pages in length.
	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.	Yes	All necessary elements of the evaluation report are included in the executive summary, and it follows the structure of the main report.
	iii Includes all significant information in a concise yet clear manner to understand the theme, intervention, programme, project and the evaluation.	Partially	While the most significant information is included, it could have been presented in a clearer manner. For example, conclusions could have been explained more clearly and broken down by key points rather than provided as narrative text. Additionally each individual recommendation should have been listed separately, to enable a clearer understanding of the evaluation, and its recommendations.
SECTION B:	BACKGROUND (weight 5%)	70%	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.	Yes	There is a clear description of the 9th Uganda CP, focusing on country priorities, and UNFPA programming within that. The country context overall is well described, as is the implementation period and geographic coverage. More detail on partners and cost / budget could usefully be provided in this section, though it is noted that some of this is provided in later sections of the report.
	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	Partially	There is a clear description of the economic and social context of the evaluand, outlining in particular the population demographics and socio-economic challenges. Key policies, strategies and frameworks within Uganda are outlined, as relevant, for example detail of sexual and reproductive trends, and government plans / policies. Gender-based violence is also described as a social and political challenge. Regional diversity in cultural, social and economic challenges are described, and the role of external assistance, particularly that of UNFPA, in both development and humanitarian spheres, outlined. However, the context section would ideally have included a description of LNOB principles, and cross-cutting issues such as disability explained in relation to the CPD and wider country context. Additionally, the context section would have added greater understanding if it been structured around the thematic areas covered by the CPD.
	iii Linkages drawn between the evaluand and the ICPD benchmarks and SDGs relevant targets and indicators.	Yes	The CPs alignment with the ICPD Programme of Action and the Sustainable Development Goals (SDGs) is assessed under EQ1. More specific detail could have been provided on relevant SDG targets and indicators, specifically as they relate the context of the CPE, particularly how SDGs 3 and 5 align with the country and CP context. It is, however, well noted that the key facts table included at the start of the report includes the status of key SDG indicators, which is an example of good practice.
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	There is an identification of key stakeholders, narratively and outlined in Annex 5 where a stakeholder map is provided. This defines and describes different types of stakeholders, including by what type of population they represent (for NGOs), for example women's rights organisations. The linkages between partners could be better explained, and there are some typos / unfilled rows in the stakeholder map, which would benefit from a proof read. Additionally the stakeholder mapping does not align stakeholders with CP outputs.
	ii Stakeholders are analysed to understand their specific rights, duties, needs, interests, concerns, and potential impact on the evaluand.	Partially	There is reference to rights holders and duty bearers, and the stakeholder map at Annex 5 is clear on which stakeholders are involved in the work of UNFPA in Uganda. Data was also collected from a wide range of stakeholders including implementing partners and beneficiaries. However, the report would benefit from a fuller analysis of the specific rights, duties and responsibilities of different stakeholders.
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 intended audience is outlined as being the UNFPA Country Office, Regional Office, Headquarters, the Executive Board, government agencies, development partners, and other stakeholders. A minor area for improvement would be to list the stakeholders and development partners who would be primarily interested, for example implementing partners.
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. Overarching objectives are to provide the UNFPA Uganda CO and partners with an independent assessment of the UNFPA Uganda 9th CP (2021-2025), as well as broaden the evidence base to inform the design of the next programme cycle. Specific objectives are outlined as being: <ul style="list-style-type: none"> • to provide an independent assessment on the relevance, coherence, effectiveness, efficiency and sustainability of the UNFPA 9th CP • to provide an assessment of UNFPA humanitarian assistance and the ability of UNFPA to connect immediate, life-saving support with long-term development objectives • to provide an assessment of the role played by the UNFPA Uganda CO in the coordination mechanisms of the United Nations Country Team (UNCT). • to identify lessons learned and formulate practical recommendations for the upcoming programme cycle. Objectives are relevant and specific. There is no reference made to changes from the ToR.
	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	The scope is clearly defined as encompassing the entirety of the 9th CP activities. Geographically this covered 40 districts in three main regions of Uganda, including those hosting refugee populations. Thematically the scope covers the 9th CP's four thematic intervention areas: (a) Sexual and Reproductive Health, (b) Adolescents and Youth, (c) Gender Equality and Women's Empowerment, and (d) Population Dynamics. While temporally the scope covers the period of the 9th CP, 2021 to 2025. It also well noted that the evaluation will include assessment of organisational and wider programmatic issues, as well as specific attention to cross-cutting issues.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	65%	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? <i>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>		

	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	<p>Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. Relevant criteria are specified and aligned with questions, in the main body of the report (Table 1) and the evaluation matrix. However, it is stated that 5 OECD DAC criteria will be applied in the evaluation, but the figure accompanying this text (Figure 1) contains 6 criteria. This may have been a minor oversight, but the diagram should have been amended to remove the "impact" criteria, which was not part of this evaluation.</p> <p>Additionally, while the narrative text states that the humanitarian-specific criteria of coverage and connectedness will also be utilised, Table 1 does not align with this, and adds a separate criterion of "coordination", as a standalone criterion. This discrepancy between text and tables / figures should be clarified, and the exact criteria (humanitarian or otherwise) applied in this evaluation should be clearly specified.</p> <p>There are also 10 separate evaluation questions, when it would have been more appropriate to have 6-8 questions and include sub-questions, where relevant, as suggested by the UNFPA Evaluation Handbook.</p>
	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.	Partially	<p>The evaluation matrix is detailed and thorough and clearly presents the evaluation criteria with each corresponding evaluation question, and lists assumptions to be tested along with indicators, data sources, and methods of analysis. This structured framework ensured that all relevant aspects of the Country Programme are thoroughly assessed, providing a clear roadmap for data collection, analysis, and synthesis.</p> <p>As noted above, the criteria of coordination has been included, and appropriate data provided for questions under this criteria. However, only one assumption was developed per evaluation question, where usually multiple would be developed to reflect the different aspects of each question, meaning they lack depth and do not support a deeper understanding of the evaluation questions. This section would also benefit from a description of how the assumptions were developed and used to test the ToC. Additionally the use and presentation of the evaluation matrix filled with collected data does not follow the standard format as per UNFPA Evaluation Handbook (see Tool 16, Annex 1 UNFPA Evaluation Handbook), which should include anonymized "informaiton sources."</p>
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.	Yes	<p>There is a clear description of the intended results of the Uganda CP, which is described narratively, and in the results framework which is included as part of the TORs at Annex 1. The annexes include two annexes which contain three reformulated ToCs (Annex 8 which contains figures of the ToCs, and Annex A which contains pdf links to the ToCs - which are not able to be opened). It is assumed that they both refer to the same ToCs, however this should be clarified and the annex which contains the broken pdf links should have been removed.</p> <p>However, the information in the ToCs at Annex 8, the narrative text and the results framework all clearly describe the interventions intended results, and parts of the results chain being tested.</p>
	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).	Partially	<p>The images of the three reformulated ToCs (At annex 8) are challenging to read in detail due to the text being very small, and some graphics overlapping with each other (particularly for the ToC relating to outcome 1). It is also assumed they contain the same information as the three broken pdf links to ToCs at Annex B, though this would benefit from clarification, along with the inclusion of a clear and easily readable figure of each ToC for each outcome. It is, however, possible to identify causal relationships and there are links between outcomes, outputs and activities. Risks and assumption are also included. Section 3.2.3 clearly outlines how the ToC was assessed, and reformulated. Additionally, while risks and assumptions are included, more comprehensive assumptions about critical factors influencing change could be added, covering all key aspects, including external factors as well as stakeholder characteristics (such as their interests, capacities and opportunities) that could affect change.</p>
	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	<p>There has been a reformulation of the three theories of change; these appear to be comprehensive and relate to three outcomes of the CP. As above the report would benefit from clarity on which annexes contain the final ToCs, and that they are presented in easy-to-read graphical format.</p> <p>While there is a general description of the process of "reformulating the ToC" in section 3.2.2. (p. 21), it lacks a description of how the "assumptions" were defined, and while the reformulated ToCs (Annexure 8, p. 169) include some "general assumptions", they are not as deep, or wide ranging, as the types of assumptions underpinning the ToC that is expected of a UNFPA evaluation, and would benefit from greater detail, including: (i) assumptions about causal relationships (causal mechanisms that need to be activated to achieve the intended results - "why" change happens); (ii) assumptions about contextual factors (external conditions that affect the achievement of the intended results - "under what circumstances" change happens); and (iii) the characteristics of stakeholders (capacities, opportunities and motivations/interests of rights holders and duty bearers that affect the achievement of the intended results - "who" makes change happen and "for whom" change will happen).</p>
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.	Yes	<p>The evaluation is theory-based, with a mixed methods approach, as well as being participatory in nature.</p> <p>Methods are clearly described and robust to meet the evaluations purpose and objectives. Qualitative data collection methods for the evaluation combined traditional approaches such as Key Informant Interviews (KIIs), Focus Group Discussions (FGDs), Document Reviews, and Observations with more innovative techniques such as Outcome Harvesting and Most Significant Change (MSC). The use of Outcome Harvesting and MSC is an excellent approach to gather as wide views as possible, and particularly to gather views from beneficiaries of UNFPA programmes, as well as partners. Field observations were also undertaken.</p> <p>Quantitative data was collected through retrospective review of records, documents, websites and online databases to obtain data on key output and outcome indicators, and is well described.</p> <p>In addition, a Health Facility Assessment for Integrated SRH, HIV and GBV Services was undertaken to determine their performance in relation to providing integrated care.</p> <p>The evaluation did not use AI.</p>
	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).	Yes	<p>Data sources are clearly described, relevant and robust. They include a good mix of qualitative and quantitative sources, with a strong focus on qualitative evidence from the voices of stakeholders and rightsholders.</p>
	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	Partially	<p>Sampling strategy for fieldwork sites is well described to ensure that core areas / activities are included, along with districts which implement both humanitarian and development work. Kampala was included primarily for stakeholder interview purposes. This is outlined narratively and in Table 2 demonstrating a wide geographic and programme range which ensure a diversity of voices.</p> <p>It is noted narratively that in the selection of participants in FGDs and interviews the evaluation team used UN Sustainable Development Group (UNSDG) programming principle of LNOB; however, there is little detail of how this was applied in practice, and how different groups of people were identified to ensure diverse perspectives were captured.</p> <p>Data collection tools for each type of KII or FGD are provided in the annexes, these are clear and appropriate, but could be more detailed in terms of gathering disaggregated data about participants.</p>

	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).	Partially	Methods are robust and particularly varied, including the assessment of health centres as a line of inquiry is particularly innovative, as well as some of the methods applied in the conduct of KIs and FGDs, such as Outcome Harvesting and Most Significant Change (MSC). Overall the diversity of data collection methods and tools should allow for rigorous testing of the theories of change and help to understand causal connections between outputs and expected outcomes. However, it is noted that MSC and outcome harvesting were not explicitly discussed as sources of evidence in the findings, it may be that there was complexity around attribution for MSC, or outcome harvesting was not undertaken with FGDs; however, if challenges existed around the use of these methods they should have been outlined in the limitations section.
	v) Clear and complete description of the methods of analysis, including explainability and full disclosure of the use of AI in the evaluation process, if applicable.	Yes	There is a clear and complete description of the methods of analysis, with a content and trend analysis of qualitative data, while quantitative data was analysed using a univariate approach to summarise key variables, as well as comparison of data across different sources to identify if targets have been met. Contribution analysis was used to assess contribution of data to elements of the logic chain being tested, and triangulation applied across different methods and data. Quality assurance processes are also well described. The evaluation did not use AI.
	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	Limitations and constraints facing the collection of data are well described, including how to mitigate against bias. However, it might have been useful to include wider limitations / constraints such as gaps in the evidence, whether disaggregated data was available, or the applicability of data collection methods across different population groups.
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	There is explicit reference to the UNEG obligations of evaluators; however, this would benefit from being contextualised for this particular evaluation, demonstrating how these obligations are applied across the evaluation process.
	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., transparency 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.	Partially	Ethical principles to which the evaluation team adhered to are outlined, and the issue of consent of participants (including parental consent for those under 18) is explicitly described. The data collection tools also contain a preamble setting out the purpose of the evaluation, however; this could be improved if how individual data would be used (and stored), confidentiality and right to withdraw were included. This section could also be improved if issues of do no harm and safeguarding were addressed explicitly, as well as including explicit reference to other ethical issues, including: obtaining informed consent from evaluation participants; protecting confidentiality; and minimizing harm, with specific detail provided on each.
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	Not rated because it was not required in the ToR.
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	Partially	Findings provide sufficient levels of evidence to systematically address all of the evaluations questions. However, an area for improvement would be in how findings are presented (as per the UNFPA Evaluation Handbook) where it is expected that the finding summaries indicate the assumptions outlined in the evaluation matrix, and identify where supporting evidence in the matrix can be found. In addition, all findings should be numbered for ease of reference.
	ii) Explicit use of the evaluation's theory of change, results chain, logical framework in the formulation of the findings.	Yes	There is explicit use of the evaluation's theory of change. Indeed, each programmatic area has an assessment of the logic framework for each of the theories of change, including consideration of assumptions and data sources, as well as links between outputs and outcomes. This is an example of good practice, as the ToC is used to assess the effectiveness criterion, which includes a comprehensive assessment of progress against the results framework and indicators included in the CPD.
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.	Yes	The evaluation uses credible forms of qualitative and quantitative data and presents both output and outcome level data, directly related to the evaluation framework. Triangulation across different types of sources is clear, for example, data from document review, interviews, FGDs, and health facility assessments are utilised. Voices of beneficiaries are used to good effect through direct quotes, though there could be better use of disaggregated data from KIs and FGDs. The report could have also included footnotes to indicate what type of stakeholder provided evidence from KIs (and for quotes) to demonstrate more clearly that a wide range of voices were included.
	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, both positive and negative, for example the challenges of consistent disability-inclusive programming were outlined, particularly regarding gaps in health worker training and the availability of assistive technologies. The success of community-based programmes in reducing incidences of GBV is well noted, and attributed in part to a synergistic approach in engaging men, duty bearers, and whole communities to reduce GBV, this clearly illustrates achievement of CP outcomes. There is a clear assessment of performance indicators as part of the effectiveness assessment for each programmatic area, and relevant comparisons and benchmarks. Greater clarity could be provided on the benchmarks for the health facility assessment, beyond the data provided by the evaluation team, which notes that they would be assessed against the scorecard and validation of the key service delivery output data from the health facilities.
	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	The evaluation makes explicit reference to the different ToCs. It assesses the ToC for each key output individually and comprehensively - for example in relation to SRHR the logic framework is assessed as being well designed and clear, but that one indicator should be dropped as there was no data to assess it, while the GEWE ToC does not include a wide enough range of indicators and could be improved. In general terms findings analyse the logical framework for key output areas from outputs to results, assessing areas where results have not been achieved to determine factors which may have influenced this. As such some causal factors are identified - for example a change to distribution systems appears to be correlated with more health facilities having a stock of contraceptives, while rates of teenage pregnancy have changed, but child marriage rates have not, partly due to cultural practices of young (or forced) marriage, and to poverty. However, the analysis would have benefited from a deeper analysis of causal factors to determine not only what occurred (or did not occur) but why. Moreover, some of the causal factors are also noted in the conclusions, but it is good practice to reflect them in the
Question 13.	Does the evaluation assess and use the intervention's Results Based Management elements?		

	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	While monitoring is mentioned in the findings sections, there is not an assessment of the CP monitoring and reporting, either by programme or theme, even though the ToR states that M&E should be considered as part of the evaluation, and it is noted under EQ 2 of the evaluation matrix, which details cross cutting issues (though primarily in relation to the monitoring of vulnerable groups reached by the CP). No amendments to this requirement are noted in the design report. Additionally M&E tools are not discussed, nor how they can/ should be used to support decision making. There is discussion of government (particularly at district level) monitoring, but this is not linked to UNFPA monitoring systems in a clear way.
SECTION F:	EVALUATION CONCLUSIONS (weight 10%)	63%	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.	Partially	Conclusions are clearly formulated and present unbiased and summative statements that respond to the evaluation questions. Conclusions are well balanced including both positive and negative aspects, and also follow the expected guidance on presentation of UNFPA conclusions. However, some conclusions are based only on one EQ, and ideally conclusions should cut across multiple EQs. In addition, most of the conclusions include elements of recommendations, noting what should be done differently, blurring the distinction between the conclusions and recommendations.
	ii Conclusions are well substantiated and derived from findings and add deeper insight and analysis beyond the findings.	Partially	Conclusions are well substantiated and derived from findings, with links to which evaluation question and criteria they relate, as well as links to recommendations. They are clearly presented by strategic or programmatic conclusions. However, they are mainly a summary of the findings section, and there is a lost opportunity to elevate the analysis through critical assessment of the findings to provide deeper analysis and insights.
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.	Yes	Lessons learned are provided as part of the findings in addressing EQ7, and explore key lesson relating to Sexual and Reproductive Health and Rights, Gender Equality and Women's Empowerment, and Population Data. As such there are clearly derived from findings, and substantiated with illustrative examples across each thematic area.
	ii Lessons learned are clearly presented and provide actionable insights on the positive aspects of the evaluand as well as any areas of improvement.	Partially	Lessons learned could be more clearly presented, at the moment they are multiple and highly varied, however they are grouped by thematic area, with actionable insights based on both positive and negative evidence, indicating areas for improvement. It would be helpful to number lessons, and in the summary section, indicate more clearly what has been summarised and why, and where overlaps in lessons lie between, and across, different thematic areas.
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	Recommendations are clearly formulated and logically derived from findings and conclusions, with a link to the conclusion to which each recommendation relates provided. However, some recommendations could be improved if there were more clearly formulated to provide clear and practical guidance for implementation.
	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 are useful, and primary users are identified. However, the utility of the recommendations could be increased if instead of stating UNFPA CO or "partners" specific teams/units in the CO were named, or specific partner organisations were named, to ensure that there is clarity on responsible actors. Deadlines could also be more clearly defined for example a time frame is given of "short term" - this would benefit from clarity as to what a short or medium time frame would be in months. Time frames should also be applied to all recommendations consistently.
	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.	No	The process by which the recommendations were developed is not well described; it is not clear if key stakeholders/ those affected by the recommendations were involved, through an ERG or other mechanism.
	iv Recommendations are clearly articulated and prioritized based on their importance, urgency, and potential impact.	Yes	Recommendations are clearly articulated and prioritised, as being of high, medium or low priority.
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	Opening pages contain all expected, and necessary information for a UNFPA evaluation. However, it is noted that only the names of the evaluators are provided, not the name of the UNFPA evaluation manager, which should be included.
	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	Annexes contain appropriate information such as ToRs, evaluation matrix, data collection tools stakeholder map, and bibliography, along with theories of change. However, the annexes themselves would benefit from a thorough proofing. As they stand they can be confusing to interpret and understand. For example, there are two annexes named Annex 1, and there are two annexes that contain the ToCs (one of which they cannot be accessed as they are included as pdfs). Annex B is blank, and annex numbers change into letters. There are also a number of proofing and formatting issues which hinder the usage and utility of the annexes to support understanding of the main report.
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 evaluation report is broadly organised in alignment with the UNFPA Evaluation Handbook (2024), with the structure of the report broken down by chapter headings. This is an example of good practice as it clearly signposts the structure of the report, as well as outlining key documents included in the annex. The report is well written and easy to navigate report with numbered sections, clear titles. It would, however, benefit from a final proof reading and formatting. One area for improvement is that the findings section could have been more clearly, and appropriately, structured.
	ii Structure and length accords to UNFPA guidelines for evaluation reports; it does not exceed number of pages that may be specified in ToR. <i>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)</i>	Yes	At 69 pages in length, the report does not exceed the expected standards for a CPE (though it should be noted that the numbering includes the executive summary, which should not be included in the main report numbering).
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 main body of the report is written in an accessible and easy to understand way for the intended audience and is generally free from grammar, spelling and punctuation errors. The annexes, however, require full proof reading and formatting.
	ii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	There is frequent use of visual aids. However, they are not all clearly (or accurately) presented. For example, figure 1 contains OECD DAC criteria, and includes one which is not part of this evaluation, while the figure of the ToCs at Annex 8 are very small and not easy to read. Reference to some annexes is confusing due to mislabelling of annexes and pdf links are included which cannot be opened.
SECTION I:	CROSS-CUTTING ISSUES (weight 10%)	50%	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 data collection methods are well designed to capture a wide range of voices and perspectives from a range of stakeholders, and tools such as outcome harvesting to enable engaged participation are included (as relevant). The evaluation also explicitly states it applied processes to ensure the LNOB approach to evaluations was used to ensure as diverse voices as possible were captured. However, this could have been strengthened if evidence and information about how many people who are marginalised or vulnerable were involved in data collection. The data collection tools also do not provide a process to capture information on the people participating in KIs or FGDs with regard to whether they are refugees, people with disabilities, members of host communities etc. If this were provided it would greatly strengthen the data collection methods.
	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	Cross cutting issues are well covered in this evaluation, with multiple evaluation questions (EQ2, EQ5) explicitly concerning cross-cutting issues, vulnerable and marginalised people and clearly reflecting the LNOB standards and principles.
	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).	Partially	Although the evaluation is alert to assessing if programmes have reached the most vulnerable, including women, young people, and specific population groups, unfortunately the data on interviewees and those data about those participating in FGDs (table 3) is not disaggregated by gender, or any other factor. While some information in the findings are disaggregated by gender, for example women accessing SRHR services, the evaluation itself does not provide further disaggregation, nor is primary data collection disaggregated, excepting where an individual is quoted. This represent a missed opportunity for this evaluation as it is clear that disaggregated data was gathered, and separate FGDs were held for men and women, but it does not appear to have been used to assess results across different groups. While disability is analysed, it is primarily to identify gaps in provision of services.
	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 evaluation.	No	An intersectional lens is not applied to the data analysis.
	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	Findings conclusions and recommendations and recommendations do address cross-cutting issues, but as noted previously this would be strengthened if disaggregated data was used more explicitly, and if an intersectional lens were applied.
	vi Inclusion of young people in the evaluation team and/or Reference Group [N/A if not requested in ToR]	Not Rated	Not applicable, as not requested in ToR.
Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators? <i>Note: this question will be rated according to UN SWAP standards with detail provided below</i>	3	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.	Partially integrated	GEEW is partly integrated into the scope of analysis, with evaluation criteria and some questions reflecting a gendered approach. There is one sub questions which specifically addresses the advancement of gender equality, while the evaluation also notes that gender will be mainstreamed across the scope of the report in its entirety as a cross-cutting issue. However, the evaluation does not assess whether sufficient information (and disaggregated data) was collected at implementation stage on specific indicators to measure gender or human rights. Additionally the integration of gender could have been more clearly considered if a single evaluation question addressed the extent to which the CP had advanced gender equality, with clear sub questions to support this.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	The methodology is partly gender responsive, with tools and analysis applied that should enable a gendered approach to data collection, such as holding separate FGDs with men and women. However the number in each FGD and disaggregation of KIs and FGDs was not provided, or analysed, even though the tools allowed for collection of this data. Additionally there was no mention of gender in the analysis methods, in terms of how a gendered analysis would be conducted or intersectionality taken into account.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	While good account is taken of GEWE throughout findings, the lack disaggregated data hinders a full analysis, and the report does not indicate where the voices of different groups are expressed. Findings, conclusions and recommendations do address gender dimensions of the CP and there are specific recommendations to address GEWE priorities. For example, recommendation 10 is to Strengthen Social Norm Change and Gender Transformative Interventions, while others relate to gendered programming areas. The evaluation would have benefitted from an intersectional approach to understand gender within the wider context of social, cultural, economic and other forms of discrimination/ marginalisation, across the analysis and formulation of findings, conclusions and recommendations.

SWAP Rating Guidance

<p>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.</p> <p>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?</p> <p>b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?</p>
<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p> <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?</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?</p> <p>d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
<p>iii The evaluation findings, conclusions and recommendations reflect a gender analysis.</p> <p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality?</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?</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described?</p> <p>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?</p>

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