UNFPA Evaluation Quality Assessment Grid

Version: May 2024

REPORT RATING SUMMARY			
Overall Rating		85%	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 (West and Central Africa) Regional Programme Formative Evaluation
Region	WCA
Country	
Year of report	2024
Business Unit/programme country *(managing evaluation)*	UNFPA West and Central Africa Regional Office
Date of assessment review (dd/mmm/yyyy)	13/12/2024
Name of assessment review firm	IOD PARC
CLASSIFICATION OF EVALUATION REPORT	
Primary SDG(s) covered *(list provided below)*	3,4, 5
UNFPA Strategic Plan areas covered *(lists provided below)*	
Three transformative results	1, 2, 3
Six outputs	1, 2, 3, 4, 5, 6
Six accelerators	1, 2, 3, 4, 5, 6
Organizational effectiveness and efficiency	Yes
Humanitarian evaluation	No
Evaluation evaluand *(e.g. country programme/intervention/policy/thematic area)*	Regional Programme
Evaluation type *(e.g. formative, summative, developmental)*	Formative
Geographic scope *(e.g. global, regional, national)*	Regional

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 a formative evaluation of the UNFPA West and Central Africa Regional Programme 2022-2025. The key strengths and weaknesses of the Evaluation Report are as follows:

Key areas of strength
• **Section B: Background** – The Evaluation Report provides a clear description of the WCA Regional programme 2022-2025 (Section 2.4, p.10-14). An area of good practice is the way in which it provides a clear explanation of the six accelerators and strategic shifts, which are central to this evaluation and are reflected in the main evaluation questions. There is a clear and relevant description of the context for the WCA Regional programme provided in Section 2.1 (p.5).
• **Section C: Evaluation Purpose, Objectives and Scope** – The purpose of the evaluation is clearly defined in Section 1.2 (p.1-2). It explains why it was needed in this point in time and its intended use.
• **Section D: Evaluation Design and Methodology** – The evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation, as described in Section 1.3 (p.2-4). The evaluation matrix is clearly presented. The data sources are clearly described.
• **Section E: Evaluation Findings** – Overall, the findings are clearly presented and are structured according to the Evaluation Matrix (Annex 6), following the structure of the evaluation main questions and sub-questions. The finding statements are highlighted in bold with supporting evidence provided for each finding. This is systematic and makes it easy to trace.
• **Section F: Evaluation Conclusions** – The conclusions are very clearly formulated. They are presented with headline conclusions supported by findings. They are drafted in a way that adds deeper insight and analysis beyond the findings as they draw from across the EQs. This can be encouraged with other evaluations to prevent siloed concluding statements.
• **Section G: Evaluation Recommendations** – The recommendations are very clearly formulated. Each recommendation is accompanied by clear supporting actions, ranging from two to six suggested actions for each recommendation. Each suggested action identifies the lead units as well as those involved in implementation. This is very clearly done and can also be highlighted as good practice.

Key areas for improvement
• **Section A: Executive Summary** – While the Executive Summary is clearly written and well-presented standalone document. It exceeds the page limit by 5 pages (it is 12 pages in total).
• **Section B: Background** – While there are areas of good practice in this section, the Evaluation Report does not provide a clear identification of key stakeholders, particularly of the key partners involved. The Inception Report mentions that the inception phase included a stakeholder mapping and analysis (p.28) but this is neither elaborated on nor included in the Annex of either reports.
• **Section D: Evaluation Design and Methodology** – The sampling strategy is outlined in Annex 7 (p.87). However, the sampling is partial as it is only provided for the selection of country-level studies (para 9-12, p.87). Similar information is provided in the Inception Report.
• **Section G: Evaluation Recommendations** – The process for developing recommendations is not described.

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.*

Recommendations for improvement
• **Section A: Executive Summary** – In order to reduce the length, evaluators are encouraged to reflect on key and necessary information, particularly in relation to the section on findings (nearly 5 pages in length). Evaluators can also consider more proportionate allocation of content between methodology, findings, conclusion and recommendation.
• **Section B: Background** – To enhance clarity and transparency, it is strongly recommended that evaluators incorporate a stakeholder map or table in the background section of the Report. This would not only aid the audience in understanding the key stakeholders involved but also provide a clearer rationale for the sampling strategy for the evaluation.
• **Section D: Evaluation Design and Methodology** – Evaluators are encouraged to clearly identify key stakeholders, either through a map or a table (as mentioned in Q3i above that a stakeholder map was not included in the evaluation and inception report), and provide the criteria used for selection, even if the process was purposive, to explain how these stakeholders were selected.
• **Section E: Evaluation Findings** – Evaluators can be encouraged to include their ratings on 'quality of evidence' based on the rubrics they developed. It is also evident that the findings section has used the indicators presented in the evaluation matrix (Annex 6) as means for assessing and analysing data for each sub-question.
• **Section G: Evaluation Recommendations** – Evaluators can consider including a few sentences to describe how key stakeholders, such as the evaluation reference group or key users, were involved in developing the recommendations.

SECTION RATINGS				
SECTION A:	**EXECUTIVE SUMMARY (weight 5%)**	50%		**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.*	No		The Executive Summary is a clearly written and well-presented standalone document that is useful for informing decision-making. Unfortunately, it exceeds the page limit by 5 pages - it is 12 pages in total (p. VIII-XIX). In order to reduce the length, evaluators are encouraged to reflect on key and necessary information, particularly in relation to the section on findings (nearly 5 pages in length). The Executive Summary also includes a half-page graphic presentation of how conclusions link to recommendations (p. XV) - this is not a requirement and is not essential for informing decision-making.
	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		The Executive Summary includes all necessary components of the evaluation report including: 1. Overview of the context and intervention (p.VIII); 2. Evaluation purpose, objectives and intended users (p.VIII-IX); 3. Scope and evaluation methodology (p.IX); 4. Findings (p.IX-XIV); 5. Conclusions (p.XIV); 6. Key recommendations (p.XV-XVI).
	iii Includes all significant information in a concise yet clear manner to understand the theme, intervention, programme, project and the evaluation.	Partially		The Executive Summary includes all significant information in a clear manner to understand the context, Strategic Plan and programme and the evaluation. While the length could be reduced, as suggested in Q1.i, there are areas that could be elaborated on. For example, under methodology, the Executive Summary says that the evaluation team carried out a *"survey in all 23 countries in the region"* but it does not say *who* it surveyed. While this is relatively minor, the evaluation team can consider more proportionate allocation of content between methodology, findings, conclusion and recommendation. In addition, the rationale for including a summary of WCA Achievements (3 pages in length), while interesting, is not clear.
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		The Evaluation Report provides a clear description of the WCA Regional programme 2022-2025 (Section 2.4, p.10-14). The geographic coverage is described in para 32, under the section on WCA context, that the region covers 23 countries including 11 LIC, 10 LMIC and 2 UMIC (p.5). In para 52, it describes that the WCA regional programme 2022-2025 was endorsed by the UNFPA Executive Board in 2021 (p.10). It also explains that this evaluation seeks to complement the internal Mid-Term Review (MTR) of UNFPA Strategic Plan 2022-2025 (para 17, p.1). Key relevant findings from the MTR are provided in Section 2.5 (p.12). In para 53, it provides the budget for the regional programme including the amount raised by the UNFPA WCA, that it surpassed its target by 16% (p.10). A description of the implementation status is provided in Section 2.4, para. 52-57. An area of good practice is that it provides a clear explanation of the six accelerators and strategic shifts, which are central to this evaluation and are reflected in the main evaluation questions. This is provided in Section 2.3, which explains what the second strategic plan seeks to achieve (para 46-48), the accelerators and strategic shifts (para 49-50), and the key 'enablers' (para. 51). A brief overview of the partners engaged in implementing the regional programme (e.g., regional multilateral bodies, main implementing partners, other UN entities, etc.) is missing.

	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 and relevant description of the context for the WCA Regional programme provided in Section 2.1 (p.5). This includes a description of the external context including recovery from COVID-19 pandemic, political instability in the region, climate and environmental change, and socio-economic diversity of the countries within the region (para 32). It describes the following in more detail: demographics (para 33), macroeconomic volatility (para 34), increasing political instability in some countries including humanitarian crisis (para 35), climate and environmental changes (para 36), effects of the COVID-19 crisis (para 37), rise in conservative values and what this means for gender equality (para 38), and shifts in donor priorities (para 39). It also includes how these contextual factors relate to the Regional programme. This section is very clearly described and can be highlighted as good practice. What can be improved is integration of LNOB perspectives including disability and inclusion. In addition, the evaluation report also provides an overview of the Status of UNFPA's Transformative Results in WCA (Section 2.2, p.7) and the UNFPA Strategic Plan 2022-2025 (p.8-10).
	iii	Linkages drawn between the evaluand and the ICPD benchmarks and SDGs relevant targets and indicators.	Yes	The linkages between the WCA Regional Programme, SDGs and ICPD benchmarks are made in para 18, which says that the Regional programmes are essential components in complementing *"country-level efforts in promoting the implementation of the Programme of Action (PoA) of the International Conference on Population and Development (ICPD) and the 2030 Agenda for Sustainable Development"* (p.1). Understanding how the strategic shifts and accelerators, as operationalised in the Regional Programme, support the achievement of ICPD PoA is a key evaluation question (p.3). Similarly, references to SDGs is made in para 22 and assessing the programme *"given the state of progress towards the SDGs"* is a part of EQ4. In addition, WCA Regional programme's targets to advance SDGs is provided in para 54. While the Evaluation Report does not provide specific SDG and ICPD targets and indicators, this is understandable given that this is an institutional evaluation.
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	The Evaluation Report does not provide a clear identification of key stakeholders, particularly of the key partners involved. The Inception Report mentions that the inception phase included a stakeholder mapping and analysis (p.28) but this is neither elaborated on nor included in the Annex of either reports, even though the stakeholder map provides the foundation for sampling for data collection. Nevertheless, Annex 5 provides a list of the individuals interviewed, that key stakeholders (internal and external to UNFPA) were interviewed. In addition, para 23 in the introduction section states mentions the key intended audiences for the evaluation, which include *"UNFPA internal structures... at the regional office,... country officers...., headquarters, and the Executive Board. Partners and stakeholders of UNFPA... UNFPA WCARO partners with international organisations like UNICEF, regional bodies such as ECOWAS and the African Union, as well as national governments and organisations."* Due to the partial information, it is rated accordingly. To enhance clarity and transparency, a stakeholder map or table should have been included in the background section of the Report or a dedicated annex. This would have aided the audience in understanding the key stakeholders involved but also provided a clearer rationale for the sampling strategy for the evaluation.
	ii	Stakeholders are analysed to understand their specific rights, duties, needs, interests, concerns, and potential impact on the evaluand.	No	This information is missing. It would be useful to include information that identifies the key stakeholders and identify what their interest are, their needs, contributions, roles and responsibilities.
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 in Section 1.2 (p.1-2). It explains why it was needed in this point in time and its intended use i.e. *"to provide evidence to support UNFPA WCARO's learning related to what works…, inform the remaining period of the current Regional plan... and informing the design of the next Strategic Plan and Regional Programme"* (para 20). It also explains that the evaluation is a milestone *"within the SP 'evaluative evidence package'"* to inform the three Strategic Plans. The intended audiences are identified in para 23 (p.2).
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 to *"assess UNFPA WCARO's organisational readiness and strategic positioning to accelerate progress towards the achievement of the 3TRs and catalyse a discussion on the design of the next Strategic Plan and Regional Program for WCA..."* (para 22, p.2). No changes have been made to the objectives from the ones included in the ToR and Inception Report.

	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 section titled 'Purpose, Objective and Scope' (p.1) provides little information on the scope of the evaluation. However, the geographic scope is found in the Background Section that the evaluation covered "all 23 countries comprising the WCA region" (para 17). It also says that the evaluation covered the period "starting in 2022 until the end of data collection in mid-year 2024" (para 17). The thematic scope is included in the Section on Evaluation Questions, which outlines what the focus of the evaluation will be on (para 25). It also highlights that performance measurements related to enhancing programme are not within the scope of the evaluation" and the justification is provided. As the evaluation report includes the key areas of scope (thematic, geographic and temporal), it is said to have met this criterion.
SECTION D:		**EVALUATION DESIGN AND METHODOLOGY (weight 20%)**	67%	**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 and sub-questions are appropriate for meeting the objectives and purpose of the evaluation, as described in Section 1.3 (p.2-4). The questions have been refined from the version provided in the ToR, to align with the SPE Inception Report and other RPEs but as relevant to the WCARO context (p.2). However, the relevant OECD DAC criteria are not specified in the Final Report (p.3-4), although this is provided for each sub-question and/or focus area in the Inception Report (p.26-26 of IR). This may be because some of the questions cut across different OECD-DAC criteria. Nevertheless, it would have been beneficial for evaluators to explain their rationale for 'dropping' the evaluation criteria from the Final Report. Another option is for evaluators to note the main corresponding criteria, alongside each question or sub-question.
	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 is clearly presented. As mentioned above, the OECD DAC criteria are not used. However, the evaluation matrix includes assessment indicators, identifies main lines of inquiry/indicator, and sources, methods and tools for each evaluation sub-question. The indicators/lines of inquiry provide a basis for analysis from which conclusions can be drawn. What can also be improved is to clearly specify stakeholders that will inform specific lines of inquiry.
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.	Not Rated	As this is a formative evaluation, it looks specifically at assessing 'organizational readiness and strategic position to accelerate progress towards achieving the three transformative results" in the WCA region (para 22 on objective of the RPE, p.2). It also explains that performance measurements related to programming are not within the scope, as the evaluation assumes that incorporating 'strategic shifts or accelerators is enhanced programming that will strengthen UNFPA's contribution to the 3TRs" (para 25, p.2).
	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).	Not Rated	As mentioned above in 7i, this evaluation is formative to assess organisational readiness and strategic position.
	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.	Not Rated	As mentioned above in 7i, this evaluation is formative to assess organisational readiness and strategic position.
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	There is a section on the methodology in the main report (p.15-16) and it is further elaborated on in Annex 7 (p.85). It describes the evaluation approach as 'inductive/adaptive… with elements of a deductive theory-based approach', how these approaches are interpreted is described in the subsequent paragraphs and in the footnotes of the page (p.85). It does not provide sufficient information to understand what elements of a theory-based approach will be used. Evaluators can be encouraged to elaborate on this aspect to enhance clarity and transparency. In addition, it is described as a participatory and utilization focused evaluation, which is elaborated on in para 6 (p.86 of Annex 7). It involved 'strong stakeholder engagement' through the evaluation at different phases including early consultations, discussion of preliminary findings and co-creation workshops. The evaluation uses mixed-methods (para 5 of p.85), including both primary and secondary data. The data sources used are elaborated below under Q8ii. The evaluation matrix is not referenced in the methodology of the main report but mentioned in the Annex under data collection methods. The use of AI in the evaluation process is not mentioned and its use is not a requirement in the ToR. However, it is referenced in the Inception Report that it will work in close coordination with the Strategic Plan Evaluation (SPE) to leverage 'evaluation tools, processes, guidelines and outputs (including AI-generated content)….' If AI was used, evaluators are encouraged to describe it here or explain why it was not used following plans during inception. Overall, while some aspects can be further elaborated on, the evaluation design and methods are relevant and robust for the evaluation's purpose, objectives and scope.
	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	The data sources are clearly described. These include a regional survey, across all 23 countries; key informant interviews at HQ, regional and country offices; focus group discussions with select country and regional representatives; country missions; and a document review of strategic documents, programmatic documents, research and evaluation reports at regional and country levels (p.15). The number of respondents/participants relevant to the primary data sources are provided in Table 5 (p.15). The data sources are further elaborated on in Annex 7 (p.86). In addition, the list of individuals interviewed for each, by role/job title, is provided in Annex 9 (p.94-97). Furthermore, Annex 10 provides a list of documents review, which includes qualitative and quantitative data (p.98).
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	Partially	The sampling strategy is outlined in Annex 7 (p.87). However, the sampling is partial as it is only provided for the selection of country-level studies (para 9-12, p.87). Similar information is provided in the Inception Report. It would have been useful to clearly identify key stakeholders, either through a map or a table (as mentioned in Q3i above that a stakeholder map was not included in the evaluation and inception report), and provide the criteria used for selection, even if the process was purposive, to explain how these stakeholders were selected. The missing description of the sampling strategy includes the regional survey (as it is unclear who the survey was shared with, although, based on the survey data collection tool provided in the Annex, each country office received one questionnaire), key informant interviews, and focus group discussions.
	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).	Not Rated	As mentioned above in 7i, this evaluation is formative to assess organisational readiness and strategic position.
	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.	Yes	There is a summary of the method of analysis in Section 3.1.3 (p.16), which is elaborated on in Annex 7 (p.88-89). It describes how qualitative and quantitative data is analysed. It also add that data was analysed by evaluation questions and were triangulated to ensure validity. It also mentions conducting a joint analysis and learning sessions with the Regional Office to strengthen the validity of findings, elicit additional findings and co-create recommendations (para 14, p.89). The evidence database (Annex 14 though provided as a separate document as it was in Excel format) was also shared to demonstrate how evidence was collated across the evaluation, based on the structure of the evaluation matrix (para 15, p.89).
	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 is a clear description of limitations and mitigating steps faced by the evaluators on p.16. This includes addressing potential gaps in evidence, such as, prioritising interviews with informants with long-standing personnel and document review, to mitigate risks of recent staff turnover or limited institutional memory. Mitigation of response bias through steps including triangulation is mentioned. An additional area for improvement is the absence of the voices of rights-holders in the evaluation, this needs to be mentioned.
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 UNEG Ethical Guidelines in Annex 12, on Ethical Principles (p.100). Contextualisation is limited to a description of the team's commitment to ensuring accountability. Evaluators are encouraged to further contextualis references to UNEG obligations of evaluators and/or principles and what this means for this evaluation.

	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 in the report.	Partially	As the evaluation did not involve rights-holders, some of the ethical considerations such as avoidance of harm are not as applicable. However, the section on ethical principles (Annex 12), describes the team's evaluation commitment to confidentiality and data protection. It also describes the team's respect for all stakeholders to ensure dignity and personal agency regardless of background. Some of these commitments, particularly confidentiality, is evident in the protocols for data collection (Annex 13). As mentioned in Q8i, it is unclear if AI was used in this report. It looks like it was used by the evaluation team of the Strategic Plan Evaluation that this team was working in collaboration with, but the evaluation report does not mention the use of AI (and in the Inception Report potential use of AI is mentioned only). One key area of concern, which is resulting in a downgrading of the score, is that the evidence database provided as Annex 14 includes the responses of individual informants. These include role titles and location that are potentially identifiable.
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	There are no specific innovative practices used in this evaluation.
SECTION E:		EVALUATION FINDINGS (weight 25%)	100%	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	The findings are clearly presented. What is worth highlighting is the progress rating that the evaluation team developed to demonstrate the extent to which the accelerators and strategic shifts are incorporated into 'relevant plans, available capacity to operationalise them,' and other indicators (p.17). There are three main ratings: nascent, developing or mature, and the criteria for each is provided (p.17). These ratings are provided at the start of each evaluation sub-question under evaluation question 1.

The findings are structured according to the Evaluation Matrix (Annex 6), following the structure of the evaluation main questions and sub-questions. The finding statements are highlighted in bold with supporting evidence provided for each finding. This is systematic and makes it easy to trace. The findings and support evidence are sufficient for addressing the evaluation questions. 17). These rating are provided at the start of each evaluation sub-question. |
	ii	Explicit use of the evaluand's theory of change, results chain, logical framework in the formulation of the findings.	Not Rated	As mentioned above in 7i, this evaluation is formative to assess organisational readiness and strategic position.
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, as seen in the evidence presented for evaluation questions including sub-questions. Triangulation is evident in the findings where different data sources are referenced such as document review, survey, or interviews with specific categories of stakeholders. In order to enhance transparency, evaluators can consider adding references as footnotes as it is not always clear whether the data source was secondary data or from interviews. It would have been useful to provide a strength or quality of evidence rubric, which they developed in the evidence database, against each Finding (Annex 14 - separate document). This is currently available in the evidence database but could have been included in the main report against each finding.
	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	The findings are both positive and negative findings statements and, in several cases, they are nuanced. Overall, the findings are clearly supported by the evidence presented. As mentioned in Q12i, evaluators could have included their ratings on 'quality of evidence' based on the rubrics they developed. It is also evident that the findings section has used the indicators presented in the evaluation matrix (Annex 6) as means for assessing and analysing data for each sub-question. It would also be beneficial for evaluators to make more explicit reference between evidence and findings, as per UNFPA Evaluation Handbook (p.82).
	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).	Yes	Although this is not an output/outcome/impact evaluation, the report examines the extent to which 'enablers' facilitated the implementation of the accelerators and strategic shifts. This is covered under EQ2 (p.42-43) and further expanded on in Annex 5 where EQ2-4 were expanded on (p.70-72).
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.	Not Rated	This was not a requirement in the ToR in order to align the Regional Programme Evaluation (RPE) with the Strategic Plan Evaluation.
SECTION F:		EVALUATION CONCLUSIONS (weight 10%)	100%	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 very clearly formulated. They are presented with headline conclusions supported by findings. While they do not directly respond to the main evaluation questions, nor are they structured as such, they do respond to the evaluation questions in terms of the extent to which the strategic shifts have been operationalised, key enablers and barriers and how UNFPA can reconceptualise the next phase/programme (EQ 1-4). The conclusion summative statements (in text boxes) are supported by evidence (subsequent paragraphs), which are derived from findings and therefore cannot be said to be biased.
	ii	Conclusions are well substantiated and derived from findings and add deeper insight and analysis beyond the findings.	Yes	As mentioned in Q14i, the each conclusion statement (presented in text boxes) is supported by evidence from the findings. Evaluators could have included the numbered evaluation questions and findings to increase the line of sight between the findings and the conclusions and also added the associated recomendations (please see p.83 of the UNFPA Evaluation Handbook). Nevertheless, the conclusions are drafted in a way that adds deeper insight and analysis beyond the findings as they draw from across the EQs. This can be encouraged with other evaluations to prevent siloed concluding statements.
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	Lessons Learned are presented in Annex 4. They are derived from findings but, as suggested above in Q14ii, evaluators could have included the numbers of the findings that the lessons draw on to increase the line of sight between the findings and lessons. Overall, the lessons are practical with some examples provided although this is not consistent for all lessons. This is rated N/A as it is not referenced or 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, this is not rated as lessons are not requested in the ToRs.
SECTION G:		**EVALUATION RECOMMENDATIONS (weight 15%)**	88%	**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 very clearly formulated. They are presented as a table with the recommendation on the far left column. Each recommendation is accompanied by the number of the conclusion or findings that it relates to, making it easy to trace the line of sight between the three sections. This is a requirement as per UNFPA Evaluation Handbook. There is also a clear figure provided in the Executive Summary to show the linkages between the recommendations and conclusions (Figure 5, p.XV). The deep dives also include recommendations but these have not been rated here.
	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.	Yes	Each recommendation is accompanied by clear supporting actions, ranging from two to six suggested actions for each recommendation. Each suggested action identifies the lead units as well as those involved in implementation. This is very clearly done and can also be highlighted as good practice.
	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 recommendations is not described in this section. Evaluators can consider including a few sentences to describe how key stakeholders, such as the evaluation reference group or key users, were involved in developing the recommendations. However, as there is some mention of co-creation workshops in Annex 7, this has been rated as 'partially' meeting the criterion.
	iv	Recommendations are clearly articulated and prioritized based on their importance, urgency, and potential impact.	Yes	The recommendations are clearly articulated with each of suggested actions given a prioritisation rating (mainly medium and high). Evaluators have also included a country categorisation to aid in prioritisation. One consideration is that evaluators could distinguish between strategic and programmatic recommendations, as per UNFPA Evaluation Handbook guidance (p.83).
SECTION H:		**REPORT STRUCTURE AND PRESENTATION (weight 5%)**	100%	**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 the name of the evaluation (p.I), the date of the report (p.I), location is specified as UNFPA WCARO (p.III), names of the evaluators (p.II), names of the UNFPA WCARO point-person (p.II), names and role/titles of the Evaluation Reference Group (p.II), Table of contents including Annexes, Tables and Figures although the page numbers for tables and figures are not provided (p.IV-V), and list of Acronyms (p.VI). It does not include the timeframe for the evaluation.
	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.	Yes	The Annexes are fairly comprehensive. It includes the core components such as the Evaluation Matrix (Annex 6), List of Respondents (Annex 9), List of sites (Annex 7 under detailed methodology), Data collection Instruments (Annex 13), and list of documentary evidence(Annex 10). The Annexes also include the Summary of Findings from Country Mission (Annex 1); Deep Dives (Annex 2 and 3); Lesson Learned (Annex 4); Detailed Findings for EQs 2-4 (Annex 5); Detailed Methodology (Annex 7); ERG ToR (Annex 11); Ethical Principles (Annex 12) and Evidence Database (Annex 14, as a separate document). The only item missing is the full Terms of Reference, what is included instead is the 'Summary Terms of Reference' (Annex 8). It would be helpful for evaluators to provide the full ToRs in the Annex.
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 with clear headings and sub-headings, making it easy to navigate.

	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 structure and length is in accordance with UNFPA guidelines. It is 52 pages in total excluding the executive summary. As it is within its limit, evaluators could have included the content from 'Detailed Findings for Sub-EQ2, EQ3 and EQ4' in Annex 5 in the main body of the report.
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.	Yes	The report is easy to understand and is written in a clear and accessible way. It is generally free from grammar, spelling and punctuation errors. Its presentation is something to highlight as good practice, particularly its table of recommendations.
	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.	Yes	The Report frequently uses infographics, tables and figures to convey key information. Overall, these are clearly presented and labelled but they are not always referenced in the text. Evaluators can be encouraged to check that all tables and figures are referenced in the text to better understand and interpret them.
SECTION I:		**CROSS-CUTTING ISSUES (weight 10%)**	76%	**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	Evaluation's data collection methods were designed to capture the perspectives of diverse stakeholders. These include: UNFPA stakeholder at a global, regional, and country-level. However, it could be strengthened by including a stakeholder map (or table) and a clear description of the stakeholder strategy to increase transparency of who were included or not included. It does not involve rights-holders, which this evaluation has less scope for as a regional strategy. It would also be good practice for evaluators to include this rationale/justification in the report.
	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 addressed cross-cutting issues, especially as they sought to understand the extent to which accelerators were used in the context, and this involved looking at human-rights based and gender-transformative approaches, focusing on 'populations left behind', emphasising 'reaching those furthest behind first' and so on.
	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	There is limited data disaggregation in the data presented in Table 5 (p.15) on respondents who participated in the programme (it includes only broad geographical locations), similarly in Annex 7 on detailed methodology. The list of persons interviewed is also not disaggregated (p. Annex 9). However, the findings section looks at implications for different population groups including gender and disability.
	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.	Yes	An intersectional lens is applied in the data analysis. This is particularly evident under Q1.2 that focuses on 'populations left behind' where one of the findings highlight the need for RP and CPs to better assess intersectionality (p.20-21).
	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.	Partially	The findings, conclusions and recommendations address cross-cutting issues, particularly focusing on leave no-one-behind. However, greater focus on disability and inclusion across all evaluation questions would ensure that cross cutting issues are more sufficiently addressed.
	vi	Inclusion of young people in the evaluation team and/or Reference Group [N/A if not requested in ToR]	Yes	The evaluation included a "young and emerging evaluator" on the team (p.II).
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*	7	**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 including the evaluation questions. This is also see in the assessment criterion on gender and human rights in the evaluation in the evaluation matrix.
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	A gender responsive methodology is not outlined in the evaluation data collection methods. Furthermore, the list of evaluation participants is not disaggregated by gender. However, it uses a diverse range of data sources and processes to ensure inclusion, accuracy and credibility. The sampling frame is not sufficiently described as mentioned in 8iii. Lastly, ethical considerations in the areas of confidentiality are described but there remains some concerns about the evidence database that includes identifiable sources (see 9ii).
	iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The evaluation findings, conclusions and recommendations reflect a gender analysis.

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