

UNFPA Evaluation Quality Assessment Grid

Version: May 2024

REPORT RATING SUMMARY		
Overall Rating	69%	Satisfactory
● ● ● ● ●	Excellent	5
● ● ● ● ●	Highly Satisfactory	4
● ● ● ● ●	Satisfactory	3
● ● ● ● ●	Fair	2
● ● ● ● ●	Unsatisfactory	1
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.		

REPORT DETAILS	
Title of the evaluation report	Formative Evaluation of the Regional Program for Latin America and the Caribbean 2022-2025
Region	LAC
Country	Multiple
Year of report	2024
Business Unit/programme country (managing evaluation)	LACRO
Date of assessment review (dd/mm/yyyy)	March 10, 2025
Name of assessment review firm	IODPARC
CLASSIFICATION OF EVALUATION REPORT	
Primary SDG(s) covered (list provided below)	3, 5, 10, 13, 16 and 17
UNFPA Strategic Plan areas covered (lists provided below)	
Three transformative results	yes
Six outputs	yes
Six accelerators	yes
Organizational effectiveness and efficiency	No
Humanitarian evaluation	Yes
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	
This evaluation report is satisfactory, with a solid contextual background, a robust design and some good analysis that by and large supports the findings. While the evaluation meets standards for methodological rigor and design, it has weaknesses in the recourse to data and evidence, reducing its overall analytical strength.	
Specific strengths include:	
<ul style="list-style-type: none"> The report has a comprehensive context description that provides good strategic, demographic, social, and economic context for UNFPA's work in LAC, comprehensively linking it to the global strategy of UNFPA. A mixed-methods evaluation approach combining document reviews, KIIs, FGDs, and an online survey at different stakeholder levels along with a multi-stage sampling framework of regional, country, and thematic perspectives is important to ensure triangulation. Some use of innovative approaches through the use of AI (though the precise usages is not described) The evaluation appropriately references UNEG evaluation ethics, ensuring confidentiality and informed consent. Findings are systematically organized according to evaluation questions. Logical and well-structured conclusions with explicit linkages to findings. Practical recommendations prioritized by short-, medium-, and long-term actions, arrived at by a consensus of stakeholders. Useful and informative annexes accompany the report. 	
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	
The report's most substantive issue is the limited provision of evidence to support sub-findings or assertions in the text, with many cases where statements are unreferenced with respect to documented primary or secondary evaluation evidence, undermining their validity. A secondary issue is the absence of the voices of community-based stakeholders (i.e. rights-holders or their representatives) which were ostensibly included in the evaluation data collection process. Other areas of improvement are as follows:	
<ul style="list-style-type: none"> The executive summary lacks depth and omits some key elements for example some conclusions and recommendations are missing or overly summarised. Country and document selection criteria are not well justified – there are gaps in how final samples were arrived at. The designation of the UNFPA Caribbean Sub-Office as a country visit is also questionable. Ethical considerations are minimally addressed, with no detailed discussion on safeguarding and AI transparency – an annex (9) purporting to cover AI safeguards is entirely missing from volume 2. There is inconsistent use of secondary data, with many statistics unreferenced, reducing credibility. Findings are often descriptive rather than analytical, lacking exploration of underlying factors or causality. Despite a variety of regional crises, humanitarian programming is weakly analyzed. There is very limited disaggregation of data underlying findings by gender, disability, and other vulnerabilities – key informants to the evaluation are not broken down by gender. Recommendations are quite generic, lacking specificity and concrete implementation guidance. There are a variety of formatting and structural issues such as inconsistent table numbering, missing references, and low-resolution graphics that detract from the professionalism of the report. 	

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.
Partially	The executive summary, presented as a standalone section on pages 12-15 is four pages, which is under the recommended range. While the section is clearly written, evaluators should strive to meet the minimum size to avoid excessive brevity and inadequate detail in these important components of the report. The UNFPA evaluation handbook has extensive guidance on drafting the executive summary of a CPE (pages ix, 80 and 92) and notes that five pages is the minimum, and a "good" summary is 5-6 pages long. Most specifically, the conclusions and recommendations are overly brief would benefit from more detail, so the document can be useful for decision-making (e.g. to understand what should be done and why).
ii	Includes all necessary components of the evaluation report, including: (1) overview of the context and intervention, (2) evaluation purpose, objectives and intended users, 3) scope and evaluation methodology, (4) summary of most significant findings, (5) main conclusions and (6) key recommendations
Partially	The executive summary includes most of the essential components, evaluation purpose and objectives, methods, findings, conclusions, and recommendations, but lacks a background/context section and a clear statement of the intended audience. The UNFPA evaluation handbook states on pages ix, 80 and 92 that the context and intended users (target audience) should be described in the executive summary. Although the context is not mentioned explicitly in the table on page 80 this is a slight inconsistency in the Handbook itself with the linked guidance sheet (page ix) which is the most comprehensive (and hence definitive) guide. The table in the Handbook should be amended to clearly include the term 'context'. Additionally, it references a country visit to "the Caribbean," which is not a country. Strictly speaking, the visit to the UNFPA Caribbean Sub-Office does not equate to a country visit in the same way as those to Brazil and the Dominican Republic, just as a visit to UNFPA Headquarters would not be considered a visit to UNFPA USA. Revising this wording for greater accuracy and consistency is recommended.

	iii	Includes all significant information in a concise yet clear manner to understand the theme, intervention, programme, project and the evaluation.	Partially	Since the summary largely reiterates the headline findings from the main report, it provides a useful overview. The same applies to the conclusions and recommendations; however, conclusions 4a and 4b are missing, and the recommendations include only the headline statements rather than a more comprehensive list of action points, which would be more useful. Given that the summary is only four pages long (below the recommended 5–7), a more thorough synthesis that accurately reflects the report's analysis is warranted.
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.	Partially	<p>Section 2.2 introduces the UNFPA global and regional response frameworks but is high-level, lacking specifics on location, partners, and budget. While annexes 2 and 3 provide useful details on populations and regional priorities, these elements are still not clearly outlined.</p> <p>The main report devotes excessive space (1.5 pages) to describing the global UNFPA Strategic Plan (SP) and aligning the evaluation with it, including a low-resolution graphic that appears to link evaluation questions to SP elements. This is followed by a discussion of LAC regional office priorities, though these largely mirror the global outputs, raising questions about the added value of this section.</p> <p>This section could have provided a more detailed and specific description of the regional programme, drawing on annex 3 while adding key activities, partners, and budget details, while reducing the focus on the global SP and MTR.</p>
	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	<p>Section 2 provides a solid background on the demographic, social, and economic context of LAC populations. However, much of the data on life expectancy, birth rates, and population growth lacks references to reliable sources, raising concerns about robustness (see paras 10-14).</p> <p>Annex 2 offers a more comprehensive contextual overview, particularly in UNFPA priority areas such as LNOB, maternal mortality, GBV, poverty, migration, and the lingering effects of COVID-19. The main section also presents an in-depth discussion of political and contextual drivers relevant to UNFPA's mandate, notably on pages 19–23, which cover SRHR/family planning and gender equality/GBV.</p> <p>However, structural issues affect the clarity of Section 2:</p> <ul style="list-style-type: none"> Some subsection headings after the political context align with the 3TRs, but adolescent pregnancy, though important, is not a TR and would fit better under Family Planning. The "Unmet Need for Family Planning" subsection begins with a discussion on aging populations and declining fertility, which does not immediately support the argument, weakening its impact. <p>Additionally, while Section 2 contains useful (though mostly unreferenced) data on gender equality and rights, it lacks substantive coverage of cross-cutting issues such as disability, vulnerability, and diversity (e.g., LGBTQI groups), which are only briefly mentioned (although more details on indigenous and afro-caribbean populations are provided in Annex 2).</p>
	iii	Linkages drawn between the evaluand and the ICPD benchmarks and SDGs relevant targets and indicators.	Yes	The report references the SDGs and the ICPD in relation to UNFPA's global strategic plan (Section 2.2, page 24), but there are no further connections made to the evaluand within the background and context sections. While a regional evaluation may not require a detailed analysis of specific SDGs, incorporating some information on the region's progress toward SDG targets, readily available online, would have enriched the context and provided a clearer linkage between UNFPA's work and broader development goals.
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	As described above, there is a comprehensive description of rights-holder stakeholders provided in section 2.1 and supported by additional details in Annex 2. Thus, the context and needs of UNFPA's constituency population is well covered. However, there is little other detail on stakeholders, specifically duty-bearers - either internal (i.e. UNFPA country offices/sub-regional offices) or external beyond a general reference to partner typologies as the audience for the evaluation. The inception report provides a generic list of stakeholders but this is not referenced in the evaluation report.
	ii	Stakeholders are analysed to understand their specific rights, duties, needs, interests, concerns, and potential impact on the evaluand.	Partially	Rights-holder stakeholders (i.e. the populations for and with which UNFPA works) are analysed with respect to the specific issues and needs that they face, specifically section 2.1 in the main report and Annex 2. However, without a stakeholder map or a detailed list of regional partners, the report lacks clarity on their roles, responsibilities, and areas of overlap with UNFPA at the regional level. Adding this information would strengthen the understanding of partnerships and collaboration. The inception report provides a generic list of stakeholders but this is not referenced in the evaluation report nor does it conform to the UNFPA evaluation handbook guidelines, which mandate inclusion of those "involved in both the preparation and the implementation of the country programme, as well as those who have been directly or indirectly affected by its implementation" and provides a template for the map.
SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)			83%	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	<p>Section 1.1 does not explicitly present or refer to the purpose of the evaluation, despite the heading including the term. Instead, the evaluators use "objective" in place of "purpose," closely mirroring the justification from the TORs without significant deviation or explanation for the terminology change. That said, the evaluation's intent is relatively clear. Clarifying the distinction between purpose and objective and ensuring alignment with the section heading would improve consistency.</p> <p>The audience for the evaluation is also presented in the section (pg. 17, para 8) with both primary and secondary users identified in a general manner (i.e. no specific business units or organizations identified beyond the LACRO and Independent Evaluation Office).</p>
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).	Partially	The criterion is rated 'partially' due to a lack of clarity on articulation of the objectives. Following the language of the TORs, the evaluation has a single primary objective (which was designated as the 'purpose' in the TOR, however) related to advancement of progress on the three TRs, and a subsidiary objective which corresponds to the single objective in the TORs. While the text follows the TOR, its organization into components does not align. Clarifying this distinction and ensuring consistency with the section heading would improve coherence. For clarity, the evaluators would have been better served by following the layout presented in the inception report, which adhered to the TOR formulation and had the purpose, objective and scope clearly titled with subheadings.

	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	Paragraphs 4-7 on pages 16-17 provide a clear and concise description of the temporal and programmatic scope of the evaluation, noting the 2.5 years of focus (2022 to mid-2024), timed to provide inputs into the remainder of the current Regional Programme cycle, the design of the subsequent one and input into the evaluation of the 2022-2025 (global) Strategic Plan evaluation. There is no explicit mention of the geographical scope of the evaluation, even to note that it encompasses all of the LAC countries where UNFPA is active, which is an omission.
SECTION D:		EVALUATION DESIGN AND METHODOLOGY (weight 20%)	70%	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.	Yes	The evaluation questions are presented in Table 1 in the methods section (3.1) and have been developed from the initial list of evaluation questions in the TOR. There have been a range of changes to the original formulation, but this is well-noted. The evaluation questions remain largely unchanged from those in the Inception Report, as acknowledged in the narrative. However, they have not been accurately transferred, resulting in formatting issues, misspellings, and inconsistencies in wording. To ensure clarity and accuracy, the table should have been directly copied and formatted from the Inception Report. The formatting issues aside (which are not part of this rating), the questions themselves appear well thought-through, with the revisions from the TOR enhancing their relevance and specificity. They are also stated to be aligned with the evaluation criteria of relevance, coherence, effectiveness and efficiency. The table is structured to indicate which criteria apply to which question. Again, however, editing/formatting issues are evident, as the column headers for the criteria have two "efficiency" criteria, and no "effectiveness", which is clearly an error not picked up from the inception report as it is present here also.
	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	The evaluation matrix, presented in Annex 4, details the evaluation questions (consistent with the Inception Report and - formatting/editing errors notwithstanding - with Table 1 in section 3). They are not presented according to the evaluation criteria (not mentioned in the matrix) but are organised according to the broad analytical headings of "future outlook", "accelerators", "strategic changes" and "facilitators" introduced in the methods section. While these headings have a logic as applied to the evaluation, it would have been useful for evaluators to explain their rationale for not including the OECD-DAC criteria. This aside, the matrix presents assumptions to be tested for each evaluation question, indicators that are logically linked to these assumptions and the evaluation questions and details sources of information for each set of assumptions (not according to indicators). One issue is that the "DPO" source of information is not explained. Another, minor, issue is that the number of the assumptions in the matrix appears to have an error - numbering jumps from A1.2, A1.2 to A4.1, A4.2 etc.
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	This is a formative evaluation with focus on improving organisational learning and accelerating progress towards the Three Transformative Results (3TRs), thus this criterion is not rated. The theories of change governing each of the three TRs are introduced in section 2, and are explained in more detail, with graphical representation of each theory of change (for each TR) presented in Annex 3. The theories of change and the integrated results and resources framework were reviewed and revised as part of the mid-term review in 2023, and as such, the evaluators were not required to revisit this exercise. This is noted in the TOR and more clearly stated in the inception report: "...the evaluation will not assess the contribution nor use a theory of change and test the associated assumptions in a model to results."
	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	This is a formative evaluation with focus on improving organisational learning and accelerating progress towards the Three Transformative Results (3TRs), thus this criterion is not rated.
	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	This is a formative evaluation with focus on improving organisational learning and accelerating progress towards the Three Transformative Results (3TRs), thus this criterion is not rated.
Question 8.		Does the report specify adequate methods for data collection, analysis, and sampling?		
	i	Evaluation design and set of methods are clearly described, and are relevant and robust for the evaluation's purpose, objectives and scope, including the use of AI in the evaluation process if applicable.	Partially	The evaluation design and methods section (3.2, pages 29-31) provides a clear, if minimally detailed, description of the methodology, noting that a mixed-methods approach was adopted, combining qualitative and quantitative data sources. The design adheres to UNEG and OECD-DAC evaluation standards and includes a participatory approach, with stakeholder engagement through semi-structured interviews, focus groups, country missions, and a 'virtual' (i.e. online) survey of country offices. However, while the methodology is broadly relevant to the evaluation's purpose and objectives, and should, if applied diligently, lead to robust findings, the brevity of the section leads to some weaknesses (discussed further below, so do not affect the rating). With respect to AI, the cover pages note that AI was used "to enhance and support content analysis in the data collection phase". However, there is no reference to AI or the specific tools used elsewhere in the report. The cover page notes the inclusion of additional details on this in Annex 9, but there is no such annex in the relevant volume - it goes to Annex 8 only.
	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	As noted above, the various tools are itemised, with both quantitative and qualitative sources noted. However, the specific quantitative sources could be more clearly described - the main tools are desk review of documentation, qualitative KIs and FGDs and a short survey of 18 country offices (the individuals targeted for completion of the survey is not specified). The survey tool may be the quantitative data source cited. While it is somewhat limited in its robustness (a small sample and use of Likert-type responses prone to positive response bias), as a complement to the other primary sources and the secondary desk review, it should add value and assist in triangulating findings.

	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>The stakeholder selection process (described on pages 30-31) includes key informants from UNFPA, partner UN agencies, national governments, civil society, and academia, ensuring institutional diversity, with the purposive sampling approach used quite typical of evaluations of this kind. The use of focus groups is important in that it targeted key populations (e.g., PWDs, those affected by humanitarian crises, Afro-descendants, feminist networks, and youth). The broad approach suggests an effort to include diverse voices, although more specific details on numbers of those representing diverse groups and voices are not provided (table 2 on page 31 should have included additional columns on FGD participants - all data should have been disaggregated by gender in this table).</p> <p>The selection criteria for the sampled field visit countries are reasonably well explained, albeit narratively, noting the justification for the three locations visited. For instance, while Brazil was chosen due to its population size and South-South cooperation relevance, the rationale for excluding other high-priority countries (e.g., those with acute humanitarian needs, notably related to the Venezuela crisis or in Haiti) is not fully explained - the methods would have been stronger with greater rationale for this. However, the process of selection (i.e. who determined the criteria and the application of them, for example the reference group) is not described.</p> <p>Another area that misses detail is the sampling process for the desk review of documentation. The evaluators note that 120 documents/publications were reviewed, but a sample of 22 were actually utilized or analyzed for the evaluation - the basis for this selection is not explained.</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).	Not Rated	As the theory of change/results framework for the LAC region is explicitly excluded from testing/revision as part of this evaluation (discussed above), this criterion is not rated.
	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.	Partially	<p>The data analysis and synthesis section (3.3) in the report has some broadly useful information on the analysis process. Most importantly, it describes the use of a systematic and structured approach to the datasets, in line with the evaluation matrix and bringing together different sources of data in the narrative analysis preparation to ensure triangulation. Thus, the logical underpinnings of the analysis are clear, and this - if applied as described - is adequate to ensure a robust report.</p> <p>However, there are three main issues. Firstly, there is some overly-complex and repetitive language describing various matrices, with insufficient detail on the cleaning and coding process for the data - in particular the qualitative data (e.g. how/who/when). Secondly, there is a reference to "Visualization of large data sets" - there do not appear to be any 'large' datasets evident from the methods description, so the point of this is not clear. Finally, as discussed above, there is no mention of the AI tools noted in the cover page.</p>
	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	<p>The risks and limitations section (3.5) is included, warranting a 'partially' rating, but could have been strengthened. It identifies only two risks/limitations, both of which may require further justification. Expanding this section with a more comprehensive description of challenges would have been useful.</p> <p>The first limitation mentions that the "complexity" of the evaluation made it difficult to capture information by level (global, regional, national, subnational). However, this is not a valid limitation, as many evaluations, both commissioned by UNFPA and other organizations, successfully span these levels, incorporating the necessary sophistication to yield robust findings. In fact, addressing multiple levels can enhance robustness by facilitating triangulation across duty-bearers and rights-holders at different levels.</p> <p>The second limitation is unclear, referring to "the loss of key informants" without specifying what this loss entails. If the evaluators are referring to staff turnover, this is a recognized risk, but one that can be mitigated. Former staff remain valid interview subjects, even if they have left their positions or the organization.</p> <p>Lastly, the mitigation strategy cited by the evaluators—that the team "proactively addressed them and triangulated data sources to ensure the success and integrity of the process"—is too generic to be meaningful. A more specific and actionable description of the mitigation efforts would have add clarity and value.</p> <p>Additional limitations - such as the comprehensiveness of the secondary dataset, the survey responses, the likelihood of bias (in the survey, documents, selection of interviewees and FGD participants) - were not addressed.</p>
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.	Yes	Section 3.4 (Ethical Considerations and Quality Control) references the UNEG principles of evaluation ethics, the UNEG Evaluation Guidelines (2020 version) and other UNEG guidance documents. These are itemised and specifically referenced, and contextualized, albeit briefly and minimally.
	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	<p>The section essentially lists the various ethical principles and standards that the evaluation was guided by, with little further contextualization or description of how they were applied specifically in the context of the data collection.</p> <p>The data tools in the annexes do not have any reference to assuring informed consent, right to withdraw etc. as is standard good practice with evaluations. The only reference to ethical considerations in the tools is an assurance of confidentiality in the online survey of UNFPA staff.</p> <p>As noted above, while AI was referenced in the second cover page, there is no further reference to it in the main report or in the annexes (Annex 9, which purports to discuss the ethical safeguards, is not in the volume of annexes).</p>
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.	Yes	The report clearly notes the use of Artificial Intelligence/Large Language Models in the "content analysis" aspect of the evaluation (noted on the first page) - this is in line with the TOR which notes an expectation that AI will be used where possible. It is positive and a potential good practice to be highlighted for other evaluations that such a practice be clearly noted from the outset and due reference made to UNFPA's policies in this regard and the relevant ethical standards provided in the annexes - although this annex (annex 9) is not in the volume of annexes. It is crucial that the report ensure full transparency in this regard by detailing how AI contributed to findings, potential biases, and how AI-generated insights were validated against human analysis. No other innovative practices were noted.
SECTION E:		EVALUATION FINDINGS (weight 25%)	63%	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	<p>Overall, the findings are presented quite clearly and systematically. They are organized according to the structure as laid out in the evaluation matrix (i.e. not per the DAC evaluation criteria - discussed above), which enables a clear and logical link between the analysis and the evaluation design - this is standard good practice.</p> <p>The levels (i.e. quantities) of evidence presented against each are largely sufficient for most questions (given the obligation to remain within a defined page limit), with a mix of primary/secondary, quantitative/qualitative evidence presented across the analysis although some evaluation questions are quite thinly represented (e.g. EQ9 has less than a half page of analysis).</p> <p>Note that the evaluation questions are presented before each findings subsection - this is good practice, but they are not clearly labelled - for example, EQs 1, 9 and 13 have no label, and other questions are labelled PE2, PE3 etc. Possibly this refers to "Pregunta Evaluacion" - if so the English translation should have been used - EQ.</p>
	ii Explicit use of the evaluand's theory of change, results chain, logical framework in the formulation of the findings.	Not Rated	As the theory of change/results framework for the LAC region is explicitly excluded from testing/revision as part of this evaluation (discussed above), this criterion is not rated. The theories of change are referenced in places in the findings, however, in relation to its relevance to some priority areas for UNFPA (e.g. adolescent pregnancy, section 4.1.1, para 75).
Question 12.	Are evaluation findings derived from credible data sources as well as a rigorous data analysis?		
	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident using multiple data sources.	Partially	<p>While the use of secondary data to support many of the findings is clear and positive, a significant issue is the lack of proper referencing for much of this data. This is problematic, as some key findings are based on unreferenced secondary evidence. For instance, in the first finding regarding the continued relevance of the 3TRs to regional needs, the evaluators present a strong argument using data on key gender/GBV and SRHR outcomes. However, much of this evidence is not properly cited, and some claims—such as the assertion in paragraph 72 (page 35) that most pregnancies among <15-year-olds result from "abuse and sexual violence"—may be questionable.</p> <p>Another example is under 4.3.4 (Data and Evidence) where the report notes issues around demands for data being linked to "questioning of the legitimacy of a Human Rights Based Approach in increasingly volatile environments and the growth of conservative groups and narratives regarding gender equality and women's rights" - this is a strong statement that should have been accompanied by some additional explanation and clear sources.</p> <p>The quality of the analysis is therefore weakened by this lack of referencing. Additionally, the voices of community members and rights-holders, who were presumably included through focus group discussions, are notably absent from the analysis, which limits the depth and breadth of the evaluation. Incorporating their perspectives would enhance the credibility and comprehensiveness of the findings.</p>
	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.	Partially	<p>The report is mixed in the degree to which the findings are supported by robust evidence. In many cases, the findings are well-supported by an analysis of secondary and primary data, with references to both (such as key informants at region or country level - though the report does not differentiate between internal and external informants, which helps with triangulation). For many findings, the evaluators make statements that are not, or not fully, corroborated or supported by clear evidence. For example, in the analysis related to megatrends (4.3.1) the section on aging notes the impact of this phenomenon on a range of countries - but there is little data to support the evidence, data which should be easily accessible. The example of Brazil in this finding is indeed supported by reference to the 2023 Brazil census, but the other countries included as examples are not similarly supported.</p> <p>In some cases, the analysis is quite difficult to parse. This is due to frequent absence of attribution of statements to clear underlying evidence. This is exemplified by an overuse of passive voice throughout the report - for example, under finding 4, para 100, the evaluators note that "the effort and development of a method for the definition of the concept of accelerator and its operationalization by the Regional Office is appreciated" - but do not specify who is making the effort and who is appreciating it. Such statements - a passive-voice claim ("it is observed", "progress has been made") with limited or no recourse to supporting evidence from primary or secondary sources - are scattered throughout the report. The report should be reviewed and such instances addressed with the underlying evidence included.</p>
	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	<p>In the main, the report does not substantively explore causal factors underlying achievements or results. They are largely presented as is, with, where present, only brief discussion of factors that drive them. There are some positive examples of key factors that lead to success - for example, under finding 7 (4.3.3 - Service Scaling), the importance of regional/global policies and initiatives to driving policy change at the national level is explored, specifically the example of UNFPA's participation in the Spotlight Initiative, although, as per the above, granular evidence that supports the analysis could be much stronger. Another positive example is around UNFPA's normative work on reproductive rights (4.3.7, see para 157) where specific UNFPA work is attributed to policy changes in favour of women's rights.</p> <p>Other causal factors are not well explained or glossed over, with potentially important examples or factors presented briefly and not unpacked adequately. For example, in finding 8 (4.3.4 - Data and evidence), the important example of UNFPA's regional partnership with the Latin American and Caribbean Demographic Centre CELADE, which underlies many of the achievements, is only briefly noted - the nature of CELADE is not even explained, even in a footnote.</p> <p>Another example is UNFPA's humanitarian work (4.3.5), which is covered in approximately 1.5 pages of analysis, which seems incommensurate with the level of humanitarian needs and response in the region. The section notes challenges around UNFPA country office capacities in terms of "humanitarian expertise, funding and supplies" but does not analyse these in any meaningful manner to determine the specific issues or the underlying drivers. Section 4.4 (Enabling Factors) does make some effort to unpack some of the drivers of programmatic achievement, but it is overly descriptive of administrative/resourcing arrangements and light on analysis and/or linking these to specific programmatic or sectoral results.</p>
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	The ToR did not require an assessment of the RBM/MEAL components of the regional programme because it was aligned with the global Strategic Plan evaluation, hence this criterion is not rated. The evaluation was, however, tasked with reviewing the data and evidence aspects of the regional programme via EQ7 ("To what extent has the Regional Programme enhanced the scalability of data and evidence for decision making?"), with the specific indicators related to this in the Evaluation Matrix noting some elements of the regional-level monitoring & reporting systems (e.g. availability of disaggregated data, reporting on relevant SDGs). The analysis under the EQ (4.3.4) exhibits very little substance related to this aspect. The section notes two examples of UNFPA knowledge products on specific data collection related to SRH issues (MILENA and MEMI), but these are not specifically related to regional MEAL systems. Notably, the availability and use of disaggregated data at regional level by the RO (per the evaluation matrix) and the SDG reporting (also in the matrix) are not discussed. There is some description of and evaluation evidence related to monitoring tools presented under EQ2 and finding 3 (page 38) which notes some recommendations made by the mid-term review to strengthen capacity in this area and also positive perceptions of the "Dashboard 2.0 of performance indicators" by LACRO and COs. The nature of 'Dashboard 2.0' is not further explained.
SECTION F:	EVALUATION CONCLUSIONS (weight 10%)	75%	Comments on Rating
Question 14:	Do the conclusions clearly present an unbiased overall assessment of the evaluand?		
	i Conclusions are clearly formulated and present unbiased summative statements that respond to the evaluation questions.	Yes	There are five/six conclusions - 4a/4b is a two-part conclusion - the rationale for this is not clear, they would have been better served as standalone conclusions. The conclusions are generally clearly formulated and explicitly aligned with the evaluation questions. They effectively summarize key findings related to programme impact, efficiency, and effectiveness, presenting structured summative statements. Importantly, they reflect evidence drawn from the findings section, ensuring consistency with the evaluation's analytical framework. Finally, the conclusions indicate to what evaluation questions that they respond and to which findings they are linked as per the prescribed practice in the UNFPA evaluation handbook.
	ii Conclusions are well substantiated and derived from findings and add deeper insight and analysis beyond the findings.	Partially	The conclusions are generally well derived from the findings - they present key takeaways from the evaluation and align them with the core analytical themes of the report. The conclusions draw directly from findings presented earlier, ensuring coherence between evidence and final statements. However, their summative nature means that the conclusions do not consistently add deeper insight or critical analysis beyond the findings. While they effectively summarize key trends and confirm the relevance of UNFPA's strategic priorities, they could provide stronger analytical reflections on challenges, trade-offs, and alternative strategic directions. For instance, while the report acknowledges the funding limitations for middle-income countries, it does not explore specific strategies to address these constraints (i.e. the 'funding' to 'funding and financing' models). Similarly, the emphasis on UNFPA's need for strong technical expertise and collaboration is presented quite generically - it is not critically examined against potential gaps in execution or missed opportunities. Strengthening the conclusions with more nuanced analysis of risks, operational barriers, and alternative strategic pathways would enhance their depth and utility for decision-making.
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 have not been specifically requested as part of this evaluation, so this criterion is not rated.
	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	Lessons learned have not been specifically requested as part of this evaluation, so this criterion is not rated.
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 five recommendations are clearly formulated and logically derived from the findings and conclusions, with each recommendation linked to one or more specific evaluation conclusions. They focus on key areas such as accelerating progress toward the Three Transformative Results (3TRs), operationalizing accelerators, improving resource mobilization, strengthening organizational capacity, and adapting the business model to regional contexts, in line with the overall structure of the report, which is positive.
	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	While the recommendations are generally well-grounded in the findings, they (including the proposed actions) are somewhat generic and non-specific, with few detailed or granular action steps or implementation guidance. There is little detail there that could not be applied to any UNFPA office, or even the organisation as a whole, or even other organisations. For example, Recommendation 3 on resource mobilisation calls for a proactive, comprehensive, and multi-level approach with diverse funding partnerships - this is sound advice, but could be made to any organisation. The recommendation, for example, does not specify which funding mechanisms should be prioritized or how to address financial constraints in middle-income countries. Similarly, Recommendation 5 on adapting the business model highlights the need for flexibility but does not critically engage with risks or trade-offs of different approaches. Strengthening the recommendations with clearer pathways for implementation and explicit references to evidence from the findings would enhance their practicality and strategic value.
	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.	Yes	The preamble text to the recommendations clearly describes their process of development, firstly in draft form by the evaluation team, then presented to the "GRE" [presumably this is a spelling mistake and "ERG" was intended] and finally with an "internal working session with LACRO" for recommendation refinement, expansion and familiarisation among key stakeholders. This is sufficient.
	iv Recommendations are clearly articulated and prioritized based on their importance, urgency, and potential impact.	Yes	Notwithstanding the generality of most of the recommendations, the language of the recommendations is clear, well-articulated and logically structured, with clear links to specific conclusions. Each recommendation is associated with relevant thematic areas, units responsible for implementation, and a priority level (short-, medium-, or long-term), which enhances their usability. The prioritisation mechanism ensures that immediate and foundational actions (e.g., advancing the resource mobilization strategy and improving operational efficiency) are placed in the short term, while more strategic shifts (e.g., evolving the business model) are designated as medium-term priorities, which is positive.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	67%	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	All of these elements are present in the opening pages with the required information in place. An area for improvement would be to add a table with information on the reference group members.

	ii	Annexes include, if not in body of report: terms of reference, evaluation matrix, list of respondents, results chain/ToC/logical framework, list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology (e.g. inception report), case study reports.	Partially	The annexes are presented in a separate volume, and include all of the specified elements here, including some supplementary data on the regional theories of change and programme context. However, there is some misnumbering of annexes in the main report, and the second cover page (disclaimer on the use of AI) refers to additional details in "Annex 9" - which is not present in the relevant volume, which includes eight annexes only.
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).	Partially	While the overall structure of the report is sound, there are a variety of formatting issues that could be amended to improve readability, navigation and the overall professionalism in appearance: - The report header text format changes from pages 4 and 5 - from page 5 it is the same format as the main text, so the eye thinks it is part of the report, rather than a header. It should be formatted differently to distinguish it clearly. - The opening pages should be in roman numerals (per UNFPA evaluation guidelines). - Acronyms not introduced at first usage of terms. - Headings/subheadings at ends of pages. - Extra spacing or line breaks. - Issues with paragraph numbering (the numbering system resets from #107 to #69 on page 42). The report would benefit from a thorough editing to address the most significant and obvious of these issues at the least.
	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	The structure is in accordance with the UNFPA evaluation handbook, and the total pages of the main report are 64 excluding cover pages (52 excluding the executive summary). There is no page limit specified in the TOR, and given that this is a regional evaluation (more in line with a CPE than an institutional evaluation) the length is within mandated limits.
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	While the report is generally well-written, there are a significant number of editing, formatting, and terminological errors. While some of this can be expected given that English is not the first language of the evaluators (or most of the audience), the errors and inconsistencies are noticeable and detract from the professionalism of the report. Some examples include: - What are called the three UNFPA "transforming" results in section 4.1.1 should be corrected to "transformATIVE" results. - Inconsistent and incorrect references to annexes - Poor use of acronyms (HRH/SHR instead of SRH or SRHR)
	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 are some, but infrequent use of visual aids in the text. While these are largely informative, there are a variety of presentation, labelling and reference issues with them, specifically: - Some illustrations are not referenced in the table of contents (E.g. illustration 1, page 24). - Some of the graphs (#s 1-3) in the table of contents are bookmarked with internal hyperlinks, some (graphs 4-8) are not, and the hyperlinks lead to the incorrect graphs. Further, the list calls them 'graphs' but the caption for #4 terms it 'chart'. - Some tables (i.e. table 1, table 4) break across multiple pages - this should be avoided wherever possible. - Table 3 and Chart/Graph 3 in the table of contents are, in fact, the same item - one of the labels should be deleted. - Table 4 in the table of contents is labelled as table 3 in the text. - Finally, some of the charts are stretched and/or quite low resolution, with fine text (e.g. chart 1) that cannot be clearly made out.
SECTION I: CROSS-CUTTING ISSUES (weight 10%)			48%	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	Although the regional nature of the evaluation (which implies more of a focus on institutional diversity of perspective - i.e. UNFPA staff, partner agencies, national and decentralized governments, civil society, and academia) presented potential challenges in drilling down to the rights-holder level, the evaluation design did include consideration of community-level representatives. Specifically, FGDs were conducted with Afro-descendant communities, feminist networks, and youth organizations, demonstrating an effort to engage with historically marginalized groups. Further, the selection of Brazil and the Dominican Republic as case study countries suggests some intent to cover migration and displacement issues, particularly due to complex border dynamics with Haiti. However, while the inclusion of Afro-descendants, youth and feminist groups is positive, there was no representation of others—such as people with disabilities, LGBTQ+ communities, rural populations, and indigenous groups in the data collection design.
	ii	Evaluation questions address cross cutting issues, such as human rights-based approach, gender equality, disability inclusion, LNOB, social and environmental standards as appropriate.	Partially	The evaluation questions (page 28) demonstrate some focus on human rights, gender equality, and the LNOB principle, with one question specifically asking how the regional programme focuses on "populations left behind" and ensuring that the most vulnerable are reached, aligning with the LNOB framework. There is also reference to megatrends such as aging, migration and climate change in terms of their integration into UNFPA regional programming. However, persons with disabilities (PWDs) are not explicitly noted, despite their relevance within the LNOB framework. The evaluation also does not directly address other intersectional vulnerabilities, such as LGBTQ+ inclusion or indigenous communities, beyond broad references to marginalized groups.
	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 some, but limited disaggregation of data in the context of vulnerability. This is primarily evident in the context section, where demographic characteristics of populations by vulnerability (e.g. the proportion/number of persons with disabilities) are presented from secondary sources. Beyond this, the evaluation report (perhaps due to the regional nature of the evaluand) does not present results data that can be disaggregated by gender, age or other vulnerability factors. There are references to women, adolescents, and, to a lesser extent, persons with disabilities (PWDs) and migrant populations, suggesting an attempt to capture diversity in program outcomes. However, the analysis lacks a systematic breakdown of differential results across these groups. While some sections mention migrants, refugees, and PWDs, there is no detailed comparative assessment of how UNFPA interventions affected these groups differently or what specific barriers they faced.

	iv	Intersectional lens is applied in the data analysis, looking at various and multiple forms of exclusion and discrimination (and how they overlap with each other) and how this may impact the performance or results of the evaluand.	Partially	Although there is reference to intersectionality in the finding related to social norms and gender (section 4.3.6), thus warranting a 'partially' rating, this is primarily with reference to a need on the part of UNFPA to generate evidence on this area. Beyond this, the findings and analysis do not provide a structured intersectional analysis, for example the exploration of multiple vulnerabilities (e.g., being a young migrant woman with a disability). The findings highlight gender and age disparities but do not explicitly assess the compounded disadvantages faced by individuals experiencing multiple layers of exclusion. Further, while there are references to specific population groups (e.g., Afro-descendants, youth, feminist organizations, and some references to migrants and PWDs), the evaluation does not provide a structured analysis of how intersecting vulnerabilities influence access to services or outcomes in the region.
	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 in the evaluation report partially address cross-cutting issues, with attention to gender equality and vulnerability, LNOB, and social/climate concerns reflected in the design and findings. The findings acknowledge marginalized groups, particularly women, adolescents, and certain vulnerable populations, emphasizing their access to health, education, and protection services. However, PWDs are not consistently analyzed, and while the LNOB principle is mentioned, there is limited systematic assessment of how different subgroups within marginalized populations were reached or impacted. Environment/climate-related issues receive good attention in the findings, with a specific subsection on megatrends including climate change, but the conclusions and recommendations make no direct reference to climate-related challenges. Additionally, disability inclusion is inconsistently addressed, with some presentation of achievements supported by the regional office on disability, but no substantive reference to disability in the conclusions and recommendation.
	vi	Inclusion of young people in the evaluation team and/or Reference Group [N/A if not requested in ToR]	Not Rated	There was no specific requirement for this in the TOR, so this criterion is not rated.
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>	4	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.	Satisfactorily integrated	The evaluation satisfactorily integrates GEEW in the evaluation objective - human rights and gender equality are mainstreamed (via the 3TRs) and thus not requiring a standalone objective. Cross-cutting issues such as gender, climate change, and disability inclusion are referenced in the evaluation scope and questions.
	ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Partially integrated	The evaluation's mixed-methods approach is designed to facilitate assessment of GEEW considerations, specifically the use of FGDs, which are structured to collect gender-disaggregated data where relevant. The qualitative component is designed to explore how program interventions impact different population groups, allowing for a deeper understanding of social norms, gender roles, and structural barriers. While the sampling strategy considers diverse stakeholders, including marginalized groups such as women and adolescent girls, it does not systematically address the inclusion of other vulnerable populations, such as persons with disabilities and LGBTQ+ individuals. That said, the evaluation matrix includes a question specifically examining how GEEW/gender transformative approaches were integrated into UNFPA programming (under EQs 3, 5, 6 & 9), indicative of effort to ensure gender considerations are embedded in the assessment and an indicator noting "special attention to disability inclusion and protecting the rights of LGBTQI communities" (EQ3). This emphasis does not, however, translate significantly to the data collection tools - the KII/FGD guides have no specific mention of gender, although the survey tool does note gender and gender transformational approaches.
	iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Partially integrated	The evaluation includes some elements of a gender analysis, but its depth and consistency vary across findings, conclusions, and recommendations. The background section acknowledges key human rights and gender equality frameworks, including references to CEDAW and LNOB, but, while the findings disaggregate some data by gender and age, this is limited and they do not consistently apply a comparative analysis to examine disparities between social groups. The conclusions and recommendations are similarly thin - there is one reference to gender in Conclusion and Recommendation 1, with a single concrete action proposed on gender equality and empowerment. It can be argued that GEEW is implied in many of the more generic, high-level recommendations and action points (but, as discussed above, this is an overall weakness of the evaluation recommendations). For example, gender equality considerations can be considered part of recommendation 2 because among the six accelerators of the UNFPA Strategic Plan are gender-transformative approaches and human rights-based approaches. These accelerators are explained in the report, albeit in a footnote (footnote 8). Similarly, conclusion 3 makes reference to the accelerators, and there is a dedicated conclusion to LNOB, which implicitly includes gender equality and human rights.

SWAP Rating Guidance

I GEEW is integrated in the evaluation scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.

- 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?
- Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?
- Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?
- 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.

- 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?
- Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?
- Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?
- Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?
- 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.

- 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?
- 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?
- Are unanticipated effects of the intervention on human rights and gender equality described?
- 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