






# GEROS Evaluation Quality Assurance Tool

Version: September 2021

## Evaluation of Adolescents Development and Participation within UNICEF Programmes

REPORT RATING SUMMARY		
<b>Overall Rating</b>	75%	Satisfactory
 Exceptional (96% - 100%)	5	
 Highly Satisfactory (87.5% - 95.99%)	4	
 <b>Satisfactory (62.5% - 87.49%)</b>	3	<b>Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence</b>
 Fair (35% - 62.49%)	2	
 Unsatisfactory (0% - 34.99%)	1	

REPORT DETAILS	
<b>Title of the evaluation report</b>	Evaluation of Adolescents Development and Participation within UNICEF Programmes
<b>Report sequence number</b>	Philippines 60-2025-19623
<b>Region</b>	EAPR
<b>Year of report</b>	2024
<b>Office</b>	Philippines Country Office
<b>Coverage (countries)</b>	Philippines
<b>ToRs present</b>	Yes
<b>Date of review (dd/mmm/yyyy)</b>	07/02/25
<b>Name of review firm</b>	IODPARC

CLASSIFICATION OF EVALUATION REPORT	
<b>Management of evaluation (Managerial control and oversight of evaluation)</b>	UNICEF managed
<b>Unicef goal areas (Alignment with strategic plan priorities)</b>	
Every child survives and thrives	No
Every child learns	Yes
Every child is protected from violence and exploitation	Yes
Every child lives in a safe and clean environment	Yes
Every child has an equitable chance in life	Yes
Gender equality (cross-cutting)	Yes
Humanitarian action (cross-cutting)	No
<b>Evaluation object</b>	Thematic area
<b>Evaluation type</b>	Summative and formative
<b>Evaluation strategy</b>	Mixed methods
<b>Evaluation design (primary method used)</b>	Theory-based
<b>Evaluation level</b>	Output & Outcome
<b>Geographic scope</b>	National
<b>Primary SDG(s) covered (number)</b>	3,4,5,6

**EQA Summary:** *The rater will provide top line issues for this evaluation relevant for feedback to senior management (positive and negative), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight best practice/added value elements and the level of complexity of the evaluation.*

### Strengths:

- The purpose and objectives of the evaluation are clearly defined, including why the evaluation was needed 'now'.
- The context of the intervention is described in detail.
- Data collection and analysis techniques and key informants are appropriate and robust.
- Ethical safeguards and obligations of evaluators are fully described. Formal ethical approval was granted to the evaluation team through HML. The evaluation employed a human-rights based approach and uses empowering language.
- The report clearly indicates the major limitations encountered in the evaluation which allows the readers to understand the limits of findings, conclusions and recommendations.
- The lessons learned are logically derived from the findings and are applicable to similar interventions.
- The process for developing and finalising the recommendations is fully described.
- The extensive participatory process with both duty bearers and rights holders is a strength of the approach, although the reviewer notes extreme difficulties that the evaluation team sometimes encountered regarding participation.

### Weaknesses:

- Findings are not systematically supported by cited evidence.
- Finding headings (e.g., evaluation questions/sub-questions) do not always match those presented in Section 1.3.4 and in the evaluation matrix (Appendix D).
- Some statements in the conclusions section lack detail and are unclearly worded. The conclusions section is largely a summary of findings with no additional insights or further analysis.
- There are recommendations that appear to be derived from one individual's point of view, hinge on unknowns/assumptions, lack required details to be actionable, and appear to contradict introductory statements.
- There are spelling and grammar errors throughout the document and a few instances of duplicative text (from a few sentences to whole paragraphs).
- Some of the visuals used are pasted into the document as low-resolution images. This prevents those using screen readers from understanding the information and the low image quality prevents readers from being able to read the text.
- The executive summary requires additional details to serve as a standalone document that could inform decision-making.
- The report is quite long which could hinder readability and uptake. It exceeds the limit set in the ToR by over 20 pages.

**Recommendations for Improvement:** *The rater will identify topline recommendations to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, resources will be cited to assist evaluation managers in overseeing future evaluations.*

To remedy some of the identified key weaknesses, the quality assurance reviewer offers the following topline recommendations for improvements:

- Revise and expand on the executive summary to serve as a standalone document that could inform decision-making. Include an overview of the intervention, evaluation objectives, intended audience and complete summary of recommendations.
- Ensure all findings are systematically supported by cited evidence (either cited within the narrative or as footnotes).
- Revise recommendations to address feedback in Question 17, below.
- Ensure each visual is high quality. Add alternate text for screen readers if images are used.
- Perform a thorough proofread and copyedit to streamline the document (i.e., decrease its length), address duplicative text, and to improve sentence/paragraph structures.

SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	50%	Comments on Rating
<b>Question 1.</b>	Can the executive summary inform decision-making?		
	i Is clearly presented, serves as a standalone document useful for informing decision making, and is of relevant conciseness and depth for key users (Maximum of 5 pages unless otherwise specified in ToR).	Partially	<p>The four-page executive summary provides a sufficient summary of the methodology and lessons learned but lacks other key information needed to function as a standalone document. Without additional details on the findings and clear conclusions, it does not provide sufficient context for decision-making or adequately justify the recommendations. Further, the summarised recommendations do not sufficiently cover the full list of recommendations in the report (e.g., the summarised version in the executive summary does not note the recommendation on adequate human, technical and financial resources or the recommendation around equity and gender analyses.</p> <p>There are also some statements in the executive summary appear contradictory. For example, under Effectiveness, the summary states that some of the ADAP strategy's targets were met halfway (e.g., ALS programmes and child-friendly government certified LGUs) and under the cross-cutting summary, the report states that after COVID, '...progress accelerated in later years to meet almost all targets as per ADAP results framework'. However, under Evaluation of ATF Platform states that '...in the end, the respective programme teams of different sections were able to achieve the results anticipated in the ADAP strategic framework though [sic] collaborative approaches.'</p>
	ii Includes all necessary elements (overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions on findings, lessons learned if requested, key recommendations) as per the ToR.	Partially	<p>The executive summary includes the evaluation's purpose, summary of the methodology employed, some key summary findings, lessons learned, and a summary of some recommendations.</p> <p>The executive summary lacks discussions around conclusions, an overview of the intervention, evaluation objectives and the intended audience.</p>
	iii Includes all significant information needed to understand the intervention and the evaluation AND does not introduce new information from what is presented in the rest of the report.	Partially	<p>While the executive summary does not introduce new information from what is presented in the rest of the report, the summary does not include all significant information needed to understand the intervention and the evaluation.</p> <p>Importantly, the executive summary does not include an overview of the intervention. This means that some summary findings are unable to be understood. For example, the executive summary does not note ADAP's three priority areas so their reference under the Effectiveness finding summary is unclear. Further, the section contains summaries of findings that lack the required level of detail to be fully understood. For example, the report states the following under Impact but does not detail what the 'lasting impacts' were: 'UNICEF's capacity-building initiatives have had lasting impacts on the Bangsamoro Youth Parliament. This success has inspired similar legislative structures across localities, creating cascading impacts and amplifying adolescent voices in policymaking.'</p>
SECTION B:	BACKGROUND (weight 5%)	57%	Comments on Rating
<b>Question 2.</b>	Is the object of the evaluation clearly described?		
	i Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Partially	While Section 1.2 provides an overview of UNICEF's adolescent programming and sufficiently describes the ADAP strategy and ATF platform, the section does not include location(s), timelines, cost/budget and implementation status.
	ii Clear and relevant description of intended rightsholders (beneficiaries) and duty bearers (state and non-state actors with responsibilities regarding the intervention) by type (i.e., institutions/organizations; communities; individuals...), by geographic location(s) (i.e., urban, rural, particular neighbourhoods, town/cites, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability . . . (as appropriate to the purpose of the evaluation).	Partially	While the report contains a description of the intended rights holders (i.e., adolescents) and duty bearers (both state and non-state actors) alongside their responsibilities/roles (Exhibit 3), the report does not provide details on the geographic locations or numbers reached with disaggregation by gender, age, diversity, etc.
<b>Question 3.</b>	Is the context of the intervention clearly described?		
	i Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Partially	<p>The report dedicates nearly nine pages to describe the context of the intervention. Section 1.2 then links the programme's implementation to the national context.</p> <p>The context section includes reference to relevant policies, socio-economic, political, cultural, power/privilege and institutional factors. The context section lightly touches upon international factors as it relates to online exploitation of children. The report would have benefited from including other international factors such as international commitments through conventions, frameworks, etc.</p>
	ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	No	The report does not include discussions on specific SDGs and relevant targets and indicators for the area being evaluated.

	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	The context section (1.1) provides sufficient detail regarding the status/needs of adolescents (i.e., rights holders/beneficiaries) in the country.
<b>Question 4.</b>		Are key stakeholders, their relationships and contributions clearly identified?		
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	Partially	Exhibit 3 (p.13) identifies the partners and stakeholders under the ADAP (including duty bearers and rights holders) and their roles. However, the linkages between them are not explored.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Yes	Exhibit 3 also identifies the roles and contributions of key stakeholders including UNICEF, other duty bearers and rights holders.
<b>SECTION C:</b>		<b>EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)</b>	<b>92%</b>	<b>Comments on Rating</b>
<b>Question 5.</b>		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 report clearly and completely details the purpose of the evaluation including why it was needed at that point in time (i.e., to inform the next Country Programme for Children), its intended use and its key intended users (Section 1.3.3 and Exhibit 4).
<b>Question 6.</b>		Are the objectives and scope of the evaluation clear and realistic?		
	i	Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	Section 1.3.1 details the three evaluation objectives and these remain unchanged from those presented in the ToR.
	ii	Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), 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 or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Partially	Section 1.3.2 provides an overview of the geographic and chronological scopes of the evaluation but lacks key information regarding the thematic scope which should have been clarified.  For example, as it relates to the thematic scope, the report states that 'The Thematic focus is on the inter-sectoral priorities and intervention strategies related to adolescents under the three main priority areas – SRH, Out-of-School Adolescents, and Adolescent Participation – as per the TOR for this evaluation (Annexure 1). However, during the inception phase, it was decided to consider results beyond these three areas which were found relevant by the evaluation team to the concept of inter-sector programming and adolescent participation.'  However, the subsection does not detail what the results beyond these three areas were. Elsewhere in the report (Section 1.2 UNICEF's adolescent programming), the authors allude to what these areas are, but the report does not clearly make this link in the scope section: 'This evaluation is primarily focused on the three stated objectives and results framework for ADAP. Additionally, it considers other outcomes pertinent to adolescent programming, even if not explicitly part of the ADAP strategy or the associated results framework. For instance, some of the outcomes related to climate change, sanitation and menstrual hygiene management (MHM) were achieved through adolescent participation and intersectoral approaches, and thus, considered in this evaluation to a limited extent.'
<b>Question 7.</b>		Is the theory of change, results chain or logic well articulated?		
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	Appendix B includes a clear description of the results and indicators that relate to this evaluation alongside baseline, target and 2023 status figures.
	ii	Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	The causal relationship between outputs and outcomes is presented in Appendix B (ADAP results framework).
	iii	For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Yes	The evaluation team provides a brief assessment of the ToC in Section 1.2 and confirms that the ToC 'appropriately acknowledged the complex and interconnected pathways between policies and governance, programmes and interventions, community participation, adolescent participation, and social behaviour change communication to achieve these goals.' The ToR did not request that the ToC/results framework be reformulated/improved by the evaluators. This assessment is further explored under the Relevance criterion in the findings section.
<b>SECTION D:</b>		<b>EVALUATION DESIGN AND METHODOLOGY (weight 20%)</b>	<b>82%</b>	<b>Comments on Rating</b>

<b>Question 8.</b>	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The 29 overall evaluation questions and 6 AFT-specific evaluation questions comprehensively cover the three evaluation objectives. Most of the evaluation questions and sub-questions align with the specified criteria, but there are some that may have been more appropriately placed. Due to space constraints in the feedback template, only two examples follow, but there were multiple instances throughout the evaluation questions and sub-questions:  Third evaluation question under Relevance: 'To what extent were the key priorities and implementation approaches used in the adolescent programming aligned with UNICEF country programme strategy and planned results?' This question is more appropriate under Coherence as it examines how well the intervention aligned with other strategies, policies and programmes within UNICEF (internal coherence). The question does not relate to the extent to which it met the needs of its beneficiaries and stakeholders (i.e., Relevance).  The second sub-question under Efficiency: 'To what extent is the results framework used for results-based monitoring?' Appears to be more closely tied to Effectiveness, not Efficiency. Without revision to this sub-question to make a link to Efficiency (e.g., if a strong results framework and results-based monitoring allowed for timely adjustments to improve resource allocation), this question should have been under Effectiveness as it assesses whether programming achieved results-based monitoring rather than how efficiently the resources were used.
	ii In addition to the questions and sub-questions, the evaluation matrix includes indicators, benchmarks, assumptions and/or other processes from which the analysis can be based and conclusions drawn.	Partially	The evaluation provides an evaluation matrix for the overall 29 evaluation questions, but does not provide a matrix for the ATF-specific evaluation questions. The evaluation team provided the following justification: 'Evaluation Matrix was not developed for the evaluation of ATF because the matrix for the ADAP strategy evaluation was expected to provide all information required to evaluate the performance of the ATF. Further, no success measures or indicators were predefined for the ATF to apply any rigorous evaluation matrix lens.'  The report would have benefited from a stronger justification as the one provided is insufficient. (E.g., explanation for why success measures or indicators were not developed by the evaluation for this evaluation).  The evaluation matrix for the 29 evaluation questions and associated sub-questions is robust and contains measures of success, data sources, data collection methods and methods of data analysis.
<b>Question 9.</b>	Does the report specify adequate methods for data collection, analysis, and sampling?		
	i Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	Section 2.2 describes the set of methods used during the evaluation. The methods are relevant and adequately robust for the evaluation's purpose, objectives and scope. Section 2.2 provides a clear summary description of the methods while Annexure 3 provides the full set of interview guides. Section 2.3 details the participatory workshops with the evaluation reference group and adolescents learning group.
	ii Data sources are appropriate - these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR) - and are all clearly described.	Yes	Data sources for primary qualitative data were appropriate and included UNICEF staff, government staff (national, provincial, city), partners (CSOs, implementing partners, other UN entities), youth networks and rights holders.  Secondary data sources were also appropriate. Secondary data analysis included existing data national-level data sets, internal monitoring data and internal audit data by UNICEF. Qualitative secondary research of programme-related documents, data and reports as well as external literature was also undertaken. The lack of quantitative data is a limitation of the programme itself and does not impact this criterion which is around the appropriateness of data sources for the evaluation. All data sources utilised by the evaluation were appropriate.

	<p>iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).</p>	Partially	<p>The sampling strategy is provided in Section 2.2.3.2 but lacks information on how diverse perspectives were captured (or, if not, why). Exhibit 8 offers some disaggregated data (male/female) for study participants, but does not do so consistently.</p> <p>The section shares that in each of the four UNICEF programme areas, two provinces were selected (one city, one rural). However, in Section 2.2.3.3 Fieldwork, the sampling strategy becomes unclear. Section 2.2.3.3 states that 'Interviews in three cities and three provinces were conducted by two teams of moderators and note takers, but with daily debrief with local research leads. In BARMM, the field work in two cities and one municipality was led by the national senior evaluators themselves.' Section 2.2.3.2 appears to indicate that for each of the four (4) UNICEF programme areas, two (2) cities were selected (a total of eight (8) cities). However, Section 2.2.3.3 references three cities and three provinces and then two cities and one municipality in BARMM. This totals five (5) cities split between three (3) provinces and one (1) municipality.</p> <p>Exhibit 8 shows that data was collected from four (4) regions and a mixture of city-level and provincial-level personnel as well as CSOs/implementing partners.</p> <p>Clarity on the sampling plan including the number of regions, provinces and cities is required. The information presented in the report does not appear to agree nor is it clearly articulated and presented.</p>
	<p>iv Clear and complete description of the methods of analysis.</p>	Yes	<p>Section 2.2, and in particular Section 2.2.3.4 Coding and Analysis clearly and completely describes the methods of analysis.</p>
	<p>v Methodology allows for drawing causal connections between outputs and expected outcomes.</p>	Yes	<p>While the mixed-methods approach and evaluation questions allow for drawing causal connections between outputs and expected outcomes, significant limitations with regard to available quantitative data hindered the possibility of confidently stating these connections. The methodology was not the issue here, rather the absence of quantitative data. As such, the evaluation fully meets this criterion.</p>
	<p>vi Clear and complete description of limitations and constraints faced by the evaluation, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).</p>	Yes	<p>Section 2.6 details some significant limitations that were out of the reasonable control of the evaluation and that could not be mitigated in some cases. Exhibit 10 details challenges that had mitigation measures.</p>
<p><b>Question 10.</b></p>	<p>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:</p>		
	<p>i Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).</p>	Yes	<p>Section 2.5.2 Obligations of Evaluators fully addresses this criterion. The section outlines how the team addressed each of the obligations (independence and conflicts of interest, impartiality, credibility and accountability) in bullet point format:</p> <p>' Independence: Evaluators affirm their independence from the programs being assessed, ensuring they have no vested interests in the evaluation outcomes. They are neither influenced by the individuals being evaluated nor by other stakeholders involved in the program. Evaluators also confirm that there are no potential conflicts of interest that could compromise the integrity of the evaluation. • Impartiality: Evaluators have documented a full range of perspectives and evidence, giving equal consideration to all relevant data, regardless of whether it supports or contradicts the program's objectives. Furthermore, the findings and interpretations have been validated by a subset of study participants who are members of the ALG and ERG. • Credibility: The evaluation methods have undergone rigorous peer review both at the time of contract awarding and during the finalization of the inception phase. The evaluation employs the most practical and effective methods to ensure results are trustworthy and credible to all stakeholders. Limitations of the evaluation are clearly listed. • Accountability: Evaluators have meticulously documented adherence to ethical procedures, ensured de-identification of data, conducted independent coding and initial analysis before consolidating findings, and performed multiple internal reviews prior to sharing the report with UNICEF.'</p>
	<p>ii Description of ethical safeguards for participants appropriate for the issues relevant to methodology and how they are applied (respect for dignity and diversity, right to self-determination, fair representation, compliance with codes for vulnerable groups, confidentiality, and avoidance of harm). For those cases where the evaluation involves interviewing children, explicit reference is made to the UNICEF procedures for Ethical Research Involving Children.</p>	Yes	<p>Section 2.5.1 Overall safeguards against risks to participants fully meets this criterion.</p>
<p><b>Question 11.</b></p>	<p>Does the evaluation incorporate innovative practice that adds value to the evaluation process?</p>		

	i Innovation practice is used to improve the quality of evaluation process. This could evident in several ways such as the design of the methodology (i.e. use of technology for data gathering, extensive participatory processes, systematic analysis processes such as collaborative outcomes reporting and incorporation of big data, specific strategies to address complexity such as outcome harvesting, strong child rights focus), or ways of sharing of evaluation results.	Partially	The evaluation featured an extensive participatory process with adolescent rights holders through the Adolescents Learning Group with two of these members also being present in the Evaluation Reference Group. The evaluation involved adolescents at key stages of the evaluation to solicit feedback on the evaluation through virtual and in-person workshops. This process is fully described in Section 2.3.
<b>SECTION E: EVALUATION FINDINGS (weight 25%)</b>		75%	Comments on Rating
<b>Question 12.</b>	Do the findings clearly address all evaluation objectives and scope?		
	i Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	The questions listed in the findings do not clearly link back to the evaluation questions in Section 1.3.4 and the evaluation questions and sub-questions in evaluation matrix in Appendix D. Because the findings section rephrases some evaluation questions and neglects to answer or highlight some sub-questions, the quality assurance reviewer can only rate this criterion as partially. Some, but not all of the questions/sub-questions are answered.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The evaluation explicitly uses the intervention's ToC and results framework in the formulation of findings. For example, Exhibit 13 within Effectiveness tracks progress towards results framework targets and key achievements/non-achievements are detailed throughout the findings section. The evaluation also assesses the ToC and concludes that it 'appropriately acknowledged the complex and interconnected pathways between policies and governance, programmes and interventions, community participation, adolescent participation, and social behaviour change communication to achieve these goals.'
<b>Question 13.</b>	Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	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 through the use of multiple data sources.	Partially	The evaluation uses credible forms of qualitative data. There were major limitations regarding quantitative data, but some secondary quantitative data were deemed credible by the evaluation team and were analysed to address some of the questions related to Effectiveness. The evaluation presents output-level data as outcome-level metrics were not included in the ADAP results framework. Unintended outcome-level results are discussed within the report based on qualitative data.  While evidence of triangulation is evident throughout the findings section, the report often lacks specific references that would a) allow readers to understand the underlying primary data (e.g., 'Supported by four women during FGDs', 'According to three government officials', etc.) and b) allows readers to know where to go for additional information as it relates to secondary data (i.e., providing the citations for secondary data). The report findings should have ensured that specific references were systematically made to supporting evidence either in the body of the report (e.g., 'Key informants from government, UNICEF and IPs confirmed that...') or as citations in footnotes. As presented, some statements clearly cite supporting evidence, but there are many examples where evidence is not cited.
	ii Findings are clearly supported by, and respond to, the evidence presented, including both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Partially	As noted above, while the report includes insightful findings and analysis, the findings themselves are not clearly supported by the evidence presented because many statements throughout the section lack reference to specific supporting evidence. The report would have benefited from systematically citing the supporting evidence throughout the findings section.  The findings present both positive and negative aspects of the programming and are rooted in the results framework's indicators and evaluation questions/sub-questions.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings also analyse the logical chain (progression -or not- from implementation to results).	Yes	Causal factors leading to the achievement, non-achievement of results or inconclusive results are clearly identified throughout the findings section. For example, factors such as untimely fund releases, government approval delays, delayed technical assistance from UNICEF, lack of funding, lack of strategy or plans to strengthen data systems to identify, target, serve and track vulnerable adolescents, the COVID-19 pandemic's resource shift and economic impact, lack of comprehensive planning for ADAP strategy in CPCB's design stage, and lack of comprehensive monitoring and adequate personnel support all hindered achievement of results. On the other hand, UNICEF support in stakeholder engagement and advocacy as well as UNICEF-supported trainings, workshops, and forums all contributed to achievement of results. The absence of quantitative data often prevented the evaluation from determining achievement levels and led to inconclusive findings. The evaluation noted and analysed available data per the results framework in Appendix B and the evaluation team developed lessons learned related to the results framework.

<b>Question 14.</b>	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Yes	<p>The adequacy of the intervention's monitoring system was discussed throughout the findings, notably in Relevance, Coherence, Efficiency, and Effectiveness. Ultimately, the evaluation determined that 'the monitoring, documentation, and review mechanisms for adolescent programming, in general, and ADAP-specific intersectoral collaboration activities have been weak.' Further, the evaluation team noted the limitations of the evaluation as it related to monitoring data: 'Use of quantitative secondary or monitoring data would have substantially added to the rigour of this evaluation. However, no relevant or useful data to aid such an analysis was available. Therefore, the evaluation team was constrained to only note this as a major limitation which needs to be addressed in the next CPC.' And that: 'Except the annual RAM reports from UNICEF, the evaluation team could not collect any monitoring reports relevant to ADAP strategy. The RAM reports also only note the numbers but not the source of such numbers which can independently verified.' Further, the report notes that the 'ATF discontinued a collaboration matrix which was meant to track its performance as a result this evaluation had to rely on self-rating and qualitative insights of ATF members.'</p> <p>Based on the evaluation's findings on the inadequacies of the monitoring system, recommendations were given to strengthen monitoring efforts.</p>
<b>SECTION F:</b>	<b>EVALUATION CONCLUSIONS &amp; LESSONS LEARNED (weight 10%)</b>	63%	Comments on Rating
<b>Question 15.</b>	Do the conclusions clearly present an objective overall assessment of the intervention?		
	i Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	Partially	As noted below, statements within the conclusions section sometime lack detail and are unclearly worded (see examples below). The conclusions section is quite brief (11 sentences of conclusions) and then displays the assessment rating per evaluation question in a table. The conclusions section require a rewrite to produce clearly formulated conclusions that summarise and add key insights and analysis per the objectives of the evaluation (e.g., organised by OECD DAC criteria).
	ii Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	<p>Overall, the conclusions are derived appropriately from the findings and the brief section highlights key strengths and limitations of the intervention. However, there are some shortcomings within the section, namely:</p> <p>Some statements do not provide enough detail. An example: 'Government involvement from the design stage has increased their program ownership and developed capacities to sustain the programmes at multiple levels.' Adding more information related to what is meant by 'sustain the programmes at multiple levels' would have led to a stronger conclusion.</p> <p>Some statements are unclearly worded. An example: 'The decisions for focusing the ADAP strategy on only three priority areas, but other relevant results areas under the larger adolescent programming concept, were not explained.'</p> <p>The section mostly restates findings rather than deeply analysing their broader implications. The section would have been strengthened if the analysis went beyond summarising the findings.</p>
<b>Question 16.</b>	Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
	i Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Yes	The lessons learned are logically derived from the findings and are applicable to similar interventions.
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	<p>Overall, the lessons learnt provide useful insights, are concisely presented and have sufficient detail. However, there are some exceptions:</p> <p>--Lesson learnt 2 requires more information to understand the lesson within the statement 'all reviewed examples indicated that these groups were held accountable for delivering results, rather than merely monitoring collaborations'.</p> <p>--Lesson learnt 3 ('Data systems are essential for rights-based programming at scale.') could be combined with lesson learnt 5 ('While scale-up is challenging, at-scale is possible but ensuring equity requires strong data and monitoring mechanisms:').</p> <p>--Lesson learnt 4 contains duplicative text.</p> <p>--Lesson learnt 7 notes 'learnings' from COVID-19, but does not provide any lessons learnt that capture these 'learnings'.</p>
<b>SECTION G:</b>	<b>RECOMMENDATIONS (weight 15%)</b>	75%	Comments on Rating
<b>Question 17.</b>	Are recommendations well grounded in the evaluation?		

	i Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Partially	The recommendations align with the evaluation purpose. Some recommendations are clearly formulated and logically derived from the findings/conclusions while others lack clarity and do not clearly link back to the findings. Examples are noted below and as feedback to QA question 17.ii, below. <b>There is a recommendation that appears to be derived from one individual's point of view</b> and the report does not confirm whether consensus was reached or how this was grounded in the findings: 'While CPC9 RFD is not signed, one of the ERG workshop participants recommended that the indicator related to budgets should not focus on the amount of allocation or increase because that is not within the direct sphere of influence by UNICEF...' <b>There is a recommendation that lacks detail to be actionable:</b> 'Support the national and/or regional governments to model/pilot ADAP as a policy across all relevant at-scale programmes (say, at the regional level).' <b>There is a recommendation that appears contradictory to the introductory statements:</b> 'Although adolescent participation is essential as per HRBA, adolescents may not have the capacity to contribute meaningfully to at stages such as during the needs assessment and programme design...UNICEF should work on a...plan such that in the next CPC there can be more effective involvement of adolescents in all stages of a programme...' <b>There is a recommendation that hinges on unknowns/assumptions:</b> 'Review RAM (assuming CPC9 RFD is a subset of it) in a multi-sectoral group workshop...'
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	Some recommendations offer actionable guidance for UNICEF, but others lack detail, rely on assumptions, use unclear or contradictory language, reflect personal opinions, or include irrelevant information. Select examples follow but this is not an exhaustive list due to space constraints in this template: <b>Unclear recommendation that lacks detail to be actionable:</b> 'Support the national and/or regional governments to model/pilot ADAP as a policy across all relevant at-scale programmes (say, at the regional level).' <b>Hinge on unknowns/assumptions:</b> 'Review RAM (assuming CPC9 RFD is a subset of it) in a multi-sectoral group workshop...' <b>Derived from one individual's point of view:</b> 'While CPC9 RFD is not signed, one of the ERG workshop participants recommended that...' <b>Appears contradictory to lead-up material:</b> 'Although adolescent participation is essential as per HRBA, adolescents may not have the capacity to contribute meaningfully to at stages such as during the needs assessment and programme design, adapting and developing tools and procedures to local contexts, and monitoring and review. Therefore, UNICEF should work on a timebound plan such that in the next CPC there can be more effective involvement of adolescents in all stages of a programme and not only during implementation.' <b>Includes irrelevant suggestions:</b> '...This should include a clear and viable sustainability or exit plan that outlines how resources will be progressively leveraged from government and other stakeholders over the implementation period...' (I.e., this sentence is more appropriate for the preceding recommendation, not this one around verifying adequate resources.)
	iii Process for developing the recommendations is described, and includes the Involvement of duty-bearers, as well as rights holders when feasible (or explanation given for why they were not involved).	Yes	The recommendations section outlines the process for developing recommendations in detail. (Note that this explanation is repeated, with the information in two paragraphs being nearly identical.) Both duty bearers and rights holders were involved in the validation and recommendation-building workshops.
<b>Question 18.</b> Are recommendations clearly presented?			
	i Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	Yes	All evaluation recommendations are directed at UNICEF. Recommendations are organised between what UNICEF could do internally and what UNICEF could do for external actors.  Each recommended action is accompanied by a prioritisation level (i.e., immediate term (<1 year), short-term (1-2 year), medium term (3-4 year), Medium-term (4-5 year), Long-term (5+ years)/Suggestions for the next CPC).
<b>SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)</b>		71%	Comments on Rating
<b>Question 19.</b> Does the evaluation report include all relevant information?			
	i Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, 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, page numbers.	Yes	The report fully meets this criterion.

	ii Annexes include: terms of reference, evaluation matrix, list of interviewees, results chain/ToC/logical framework (unless included in report body), 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, information about the evaluator(s), etc.	Yes	All of the information required in this criterion is presented in the body of the report (e.g., primary data participants by stakeholder type and location-Exhibit 8) or as annexes.
<b>Question 20.</b>	Is the report logically structured?		
	i Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	The structure is easy to identify and navigate with numbered sections and subsections with clear titles and sub-titles. The Table of contents are linked for easy access as well. Overall, the report is a well-formatted document. One exception is that Exhibit 7. Overview of Evaluation Methods in Section 2.2.1 should have followed the first paragraph of Section 2.2 Evaluation methods. Since the exhibit is referenced in the text of Section 2.2, paragraph 1, it does not impede the reader's ability to absorb the information, but should not have been presented under Section 2.2.1 Desk review.
	ii Structure accords to UNICEF guidelines for evaluation reports: context, purpose and methodology would normally precede findings, which would normally be followed by conclusions, lessons learned and recommendations.	Partially	The structure largely follows UNICEF guidelines for evaluation reports. However, the report presents conclusions within the 'Evaluation Findings' section. Findings and conclusions should be separate: the findings section should present the objective results and data from the evaluation, while the conclusions section should provide interpretations and insights drawn from those findings.
<b>Question 21.</b>	Is the report well presented?		
	i Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	No	The ToR states that the final report should ideally be 40 pages but not to exceed 70 pages. The final report is 92 pages, exclusive of the executive summary and annexes. Note that there are some duplicative text within the report (e.g., duplication of bulleted list on pp. 38-39, duplication of the process for developing recommendations, duplication of language in lesson learnt 4 on page 82). The report would have benefited from thorough proofreading and copyediting to eliminate redundant text and enhance conciseness throughout.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	The report is written in an accessible manner that is easy to understand. The report is generally free from grammar, spelling and punctuation errors. The errors that are present (e.g., COVID-10 vs COVID-19, though vs through, revie vs review, ADP vs ADAP and unclear sentences like 'Therefore, the local programming responded to what was needed and critical at that time but used and used the principle of ADAP strategy.') do not impact a reader's ability to understand the content.
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labeled, and referenced in text.	Partially	The report frequently uses visual aids including graphics, figures and tables (all titled as 'exhibits') to convey key information. All labeled exhibits are clearly presented, labeled and referred to in the text. The graphics are not always easy to read due to poor image quality (E.g., Appendix A Theory of Change for Adolescent Programming). Further, some visuals are provided as pictures and do not contain alternate text which prevents those using screen readers from accessing the information.  There are also some instances of text boxes that do not contain the proper labels or references in the list of figures/exhibits (e.g., the four case studies on pages 48-49 and the tables within the recommendations section).
<b>SECTION I:</b>	<b>EVALUATION PRINCIPLES (weight 10%)</b>	<b>88%</b>	<b>Comments on Rating</b>
<b>Question 22.</b>	Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?		
	i Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	Section 2.1 Evaluation Approach references a human rights-based approach and confirms that the evaluation utilised a human rights-based framework which supports UNICEF's 'commitment to addressing systemic inequalities and promoting sustainable, rights-driven development outcomes as per the Convention on the Rights of the Child (CRC), the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW), and other global rights frameworks.'
	ii Clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation (for example, a reference group is established, stakeholders are involved as informants or in data gathering).	Yes	Section 2.3 details the level of participation of key rights holders as members of the adolescents working group and key duty bearers as members of the evaluation reference group. The evaluation conducted multiple participatory workshops with both of these groups throughout the evaluation. This gave both duty bearers and rights holders the opportunity to provide feedback on the evaluation methods and tools, findings and recommendations. Section 2.2 details the level of participation of rights holders and duty bearers as informants (see Exhibit 8).

	<p>iii Language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias, among others; use terminology of rights holders and duty bearers; data is disaggregated by marginalized group; differential results are assessed (distribution of results across different groups).</p>	Partially	<p>The report does an excellent job at using language that is empowering and inclusive (e.g., avoids gender, heterosexual, age, cultural and religious biases). The terms rights holders and duty bearers are used throughout, with only a few references to 'beneficiaries'. Due to quantitative data limitations, many data points were not disaggregated by marginalized groups and differential results were not assessed, with some exceptions (e.g., DHS indicators with information on young girls (15-24 years) and senior women (30-39 years)).</p> <p>Some data gathered by the evaluation team were disaggregated by sex, but not all (see Exhibit 8 where the sample is sometimes disaggregated by sex at the sub-national level but not consistently, and no disaggregation is provided for national-level participants).</p>
	<p>iv Evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind (gender and other excluded and marginalized groups). It is disability inclusive.</p>	Yes	<p>Questions and associated findings related to gender and equity (cross-cutting themes but presented in Section 3.8) specifically sought to evaluate the extent to which the programme addressed child rights and Leave No-one Behind. Issues around inclusion of marginalized groups (e.g., persons with disabilities, diverse gender identities) were also explored through other evaluation criteria such as Relevance, Coherence and Effectiveness. The evaluation findings are disability inclusive as they included an assessment of the extent to which the programme addressed issues related to adolescents with disabilities (to the extent that data allowed). As such, the report fully meets this criterion. However, the report would have been stronger with a summary of how the evaluation methodology ensured disability inclusion.</p>
<b>Question 23.</b>	Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)	8	
	<p>i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.</p>	Fully integrated	<p>a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results? YES  b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? MAINSTREAMED  c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? STANDALONE CRITERION AND MAINSTREAMED  d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? YES</p>
	<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p>	Satisfactorily integrated	<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? PARTIALLY  b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? YES  c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? YES  d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? YES  e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? YES</p>
	<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p>	Fully integrated	<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? YES  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? YES  c. Are unanticipated effects of the intervention on human rights and gender equality described? YES  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? YES</p>

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

- 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?