

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Formative Evaluation of Gender Transformative Programming through Investment in Adolescent Girls' Leadership in Eastern and Southern Africa (ESAR)

REPORT RATING SUMMARY			
Overall Rating	77%	Satisfactory	
●●●●●	Exceptional (96% - 100%)	5	
●●●●○	Highly Satisfactory (87.5% - 95.99%)	4	
●●●○●	Satisfactory (62.5% - 87.49%)	3	Meets UNICEF/UNEG standards for evaluation reports and decision makers may use the evaluation with confidence
●●○●●	Fair (35% - 62.49%)	2	
●○○●●	Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS			
Title of the evaluation report	Formative Evaluation of Gender Transformative Programming through Investment in Adolescent Girls' Leadership in Eastern and Southern Africa (ESAR)		
Report sequence number	ESARO, Kenya/63/2024/18927		
Region	ESAR		
Year of report	2023		
Office	UNICEF ESARO		
Coverage (countries)	Malawi, Mozambique, Namibia, Tanzania, Uganda, Zimbabwe		
ToRs present	Yes		
Date of review (dd/mm/yyyy)	July 2, 2024		
Name of review firm	IOD PARC		
CLASSIFICATION OF EVALUATION REPORT			
Management of evaluation (Managerial control and oversight of evaluation)	UNICEF managed		
Unicef goal areas (Alignment with strategic plan priorities)			
Every child survives and thrives	Yes		
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)	Yes		
Evaluation object	Programme		
Evaluation type	Formative		
Evaluation strategy	Mixed methods		
Evaluation design (primary method used)	Participatory		
Evaluation level	Outcome		
Geographic scope	Multi-country		
Primary SDG(s) covered (number)	1, 4, 5		
EQA Summary:	<i>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.</i>		
<p>The report has a number of strengths:</p> <ul style="list-style-type: none"> •The Executive Summary is clear and it serves well as a standalone document useful for informing decision making. It is of relevant conciseness and depth for key users, presenting all main findings, conclusions and summary recommendations in a clear and tidy manner. •There is a clear and relevant description of the intervention's context. This includes a clear and relevant description of the status and needs of the rightsholders of the intervention. •The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope. They are fully and clearly described. •The evaluation integrates gender aspects well and it specifies how gender issues are addressed in the methodology appropriate to evaluating GEWE considerations. A diverse range of data sources and processes were employed to guarantee inclusion, accuracy and credibility. •The findings are presented with appropriate evidence to back them up. •Ethical principles are clearly presented and adhered to. •The conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. •The recommendations align with the evaluation purpose. They are clearly formulated and derive logically from the evaluation findings and they are useful and actionable for the primary intended users and uses. <p>The report had some weaknesses, however:</p> <ul style="list-style-type: none"> •The lessons are not clearly and concisely presented; they are a mix of findings and recommendations and are not formulated as lessons. •There are some shortcomings in the report's formatting, in that it lacks a table of contents for the boxes, and some annexes are missing. •The report far exceeds the page limit agreed upon in the inception phase. 			
<p>Recommendations for Improvement: <i>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.</i></p> <p>The following are recommendations for improvement:</p> <ul style="list-style-type: none"> •The report would benefit from presenting the lessons in a clear and concise manner. The difference between Facts, Findings, Lessons and Recommendations have been defined by IEG (the World Bank's Independent Evaluation Group) as follows and might be useful: The Difference Between Facts, Findings, Lessons, and Recommendations Fact: What happened—an event and data (results). Not in dispute. <i>Example:</i> "The project manager was dismissed in year 5." Finding: What the analyst interpreted or concluded from the facts specific to the project. Can be disputed. □ <i>Example:</i> "Mainly because replacement of the project manager was delayed, the project did not meet its targets." Lesson: The broader significance of a finding. It draws a conclusion from experience that may be applicable beyond the project under review. □ <i>Example:</i> "Poor performance by project managers can critically affect project outcomes." Or: "If skilled project managers are not in place, it may affect the project performance negatively". •Recommendation: Suggests how to proceed in the future in the light of this experience. Proposes actions. □ <i>Example:</i> "The borrower should ensure that key project management positions are filled with competent staff. The Bank should help ensure this through appropriate covenants and prompt supervision." •The report would benefit from better formatting prior to submission, and for the text to be edited and shortened down to the agreed upon maximum pages. 			
SECTION RATINGS			

SECTION A:	EXECUTIVE SUMMARY (weight 5%)	100%	Comments on Rating
Question 1.	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).	Yes	The executive summary is clearly presented, it serves as a standalone document useful for informing decision making. It is of relevant conciseness and depth for key users. It is 6 pages, keeping within the agreed length as stated in the Inception Report to be 5-7 pages.
	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.	Yes	The executive summary includes all the necessary elements such as an overview of the intervention, evaluation purpose, objectives and intended audience, evaluation methodology, key conclusions and findings, lessons learned and key recommendations, as per the ToRs.
	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.	Yes	The executive summary includes all significant information needed to understand the intervention and the evaluation and it does not introduce new information from what is presented in the main body of the report.
SECTION B:	BACKGROUND (weight 5%)	86%	Comments on Rating
Question 2.	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.	Yes	This evaluation covers UNICEF's work with Gender Transformative Programming in the East and Southern Africa Region (ESAR), where UNICEF's work is led by the UNICEF ESARO and Country Offices (COs) and UNICEF is active in 21 countries across the region: Angola, Botswana, Burundi, Comoros, Eritrea, Eswatini, Ethiopia, Kenya, Lesotho, Madagascar, Malawi, Mozambique, Namibia, Rwanda, Somalia, South Africa, South Sudan, Republic of Tanzania, Uganda, Zambia and Zimbabwe. Six of these countries were chosen for case studies; Mozambique, Namibia and Uganda were selected for in-country case studies, while Malawi, Tanzania and Zimbabwe were selected for desk based country case studies for this evaluation. Given the diverse economic, political, cultural, and social contexts across the 21 countries in this evaluation, UNICEF's work in the region varies significantly. Providing detailed descriptions of the locations, timelines, budgets, and implementation status for all interventions is neither feasible nor necessary for this evaluation. Table 1 on page 13 provides a good overview of key markers of UNICEF's work and the context across the 21 countries, however.
	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).	Yes	The report provides a clear and relevant description of the intended rights holders and duty bearers. However, due to the nature and scope of the evaluation (thematic and regional), all duty bearers and rights holders are not presented by type, geographic location and in terms of numbers reached, etc., this is considered appropriate to the purpose of the evaluation.
Question 3.	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.	Yes	There is a clear and relevant description of the regional context of the intervention, by presenting the status of gender equality in the region, supported by statistics and other examples. It clearly shows how the implementation of the regional intervention relates to the context (pages 11-12).
	ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	SDGs 1, 4 and 5 are mentioned (page 10) as SDGs adolescent girls are "key to accelerating progress against", but no clear links are drawn from the topic/program subject to this evaluation to relevant SDGs, nor to their relevant targets or indicators.
	iii Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	There is a clear and relevant description of the status and needs of adolescent girls in East and Southern Africa, supported by statistics and other well founded examples (page 11).
Question 4.	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).	Yes	Due to the nature and scope of the evaluation (thematic, regional), a detailed identification of implementing agencies, development partners, rights holders and additional duty bearers and the linkages between them would not be possible nor relevant. There is sufficient information about rights holders and duty bearers of relevance in the report.
	ii Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	There is no identification of the specific contributions and roles of key stakeholders (financial or otherwise) in the report. This is, however, not central to the focus of the evaluation. There is a good overview of the budget of UNICEF's previous and current CDPs in the 21 countries in table 1 on page 13.
SECTION C:	EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	40%	Comments on Rating
Question 5.	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.	Partially	It is not clearly stated what the evaluation's purpose is (i.e. the word "purpose" is not used), but the report states that the evaluation "aimed to identify learning and good practice in relation to how current programming, investment and practices in adolescent girls' leadership, voice and agency, are shaping UNICEF's gender transformative programme approach in ESA." The evaluation's intended use is clear in that it will produce dissemination products to support the sharing of learning and recommendations emerging from the findings for all the different target audiences, including an adolescent friendly product that can be shared widely by the Regional Adolescent Advisory Group and adolescent panel members. It is expected that the evaluation will provide useful lessons to guide future programming, and advocacy to strengthen support and funding for adolescent girl-focused transformative programming. The evaluation team also notes that a global evaluation of the GAP 3 is planned to start early 2024, and it is hoped that this regional evaluation will provide relevant reflections and learning for that work and the key intended users of the evaluation includes the UNICEF ESARO Gender Team, ESARO Gender and Adolescent Working Group, Regional Office management and other sectors as well as the COs in the region. Other UN staff, global, regional and national partners, and donors will also be interested in the findings and recommendations, as will practitioners working with gender and adolescent issues.
Question 6.	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).	Partially	There is a description of what the evaluation seeks to achieve by the end of the process, as illustrated in Figure 3, page 17, but it is not very clear. It is stated in the introductory text to the figure that the final element in the figure is the anticipated outcome of this evaluation. However, the contents of those boxes is a mix of potential actions by UNICEF ESARO and evidence emerging from the evaluation, - and further potential actions by other actors. This is not a clear or complete description of what the evaluation seeks to achieve. The evaluation's objectives are clearly stated in the ToRs and in the report, however: a) Provide a preliminary assessment of UNICEF's GTP in ESAR vis-à-vis issues of relevance, coherence, effectiveness, efficiency, sustainability and impact with a specific focus on programming investing in girls' voice, agency and leadership; b) Examine if the interventions currently being implemented or in the pipeline are evaluable, determining whether or not the programmes are adequately defined and their results verifiable; c) Identify key lessons and the most impactful approaches, unpacking what kinds of investments have been made so far and identifying the structure and resources necessary for realising UNICEF GAP 3 (2022-2025) commitments; and d) Generate actionable recommendations that enable country offices (COs) to support replication and scale-up of more ambitious gender transformative interventions across sectors and the region with a specific focus on adolescent girls' leadership.
	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	There is a clear and relevant description of the scope of the evaluation regarding what will be covered, geographically, thematically and in terms of interventions. There are no specified limits however, or descriptions/specifications of what will not be covered.
Question 7.	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.	Partially	There is no clear description of the intervention's intended results beyond the implicit objective of increasing girl adolescents' voice, agency and leadership in the main text of the report. The results matrix in Annex 2, however, provides a clear and specific introductory sentence outlining the hypothesis of the intervention: "Investing in adolescent girls' agency, voice and leadership, as a key aspect of gender transformative programming, will lead to greater chance of impact across programmes and sectors."
	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).	No	The causal relationship between outputs and outcomes is not presented in narrative or graphic form. The "if x, then y" - figure presented on page 17 does not divide between outputs and outcomes.
	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.	Not Rated	This is not rated, as the evaluation is participatory and not theory based. The evaluation provided a theory of change (Figure 3, page 17), but this is a weak ToC only with one level of "if x, then y" statements. It furthermore presented a ToC for the evaluation itself and not for the intervention it evaluated. Furthermore, the "then"-boxes contain mostly description of further activities and not programme results. The figure does not provide the logical links (as would be expected from a logical chain/theory of change) between inputs/activities and its expected outputs and the link between the outputs and expected outcomes, neither for the evaluation, nor for the intervention.
SECTION D:	EVALUATION DESIGN AND METHODOLOGY (weight 20%)	82%	Comments on Rating

Question 8.	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	Most of the evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. However, the purpose of the evaluation itself is only vaguely defined and the main evaluation question (To what extent are UNICEF ESAR COs investing in adolescent girls' voice, agency and leadership as a central approach within this gender transformative programming?) mainly seeks to assess to what degree an approach is applied within an intervention process.
	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.	Yes	In addition to the questions and sub-questions, the evaluation matrix includes assumptions, indicators, sources of information and methods/tools for data collection.
Question 9.	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	The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope, and they are fully and clearly described on pages 19-26.
	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	The data sources are appropriate. Data collection consisted of key informant interviews and focus group discussions (73 KIIs and FGDs), data collection with adolescent girls and boys (112 young people), document review (300 documents), and a regional survey (166 respondents).
	iii Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	The sampling strategy is provided, both for key stakeholders and for the case studies, and it includes a description of how diverse perspectives were captured.
	iv Clear and complete description of the methods of analysis.	Partially	There is a clear, if not complete, description of the methods of analysis on page 28, where it is stated that "The evaluation team used thematic analysis approach to analyse qualitative data from KIIs and FGDs and to identify key trends and patterns across the evaluation questions. A coding framework was designed in alignment with the evaluation questions and assumptions outlined in the matrix and key indicators in the Gender Equality and Social Inclusion (GESI) Continuum. Emerging evidence was mapped against the GESI continuum." How the analysis of the survey was done is not described.
	v Methodology allows for drawing causal connections between outputs and expected outcomes.	No	The methodology (participatory, with a weak ToC without specified output and outcome levels) does not allow for drawing causal connections between outputs and expected outcomes. The evaluation report does not clearly divide between the two levels of results nor does it clearly report on the causal connection between them.
	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).	Yes	The key limitations and constraints faced by the evaluation are outlined in detail in table 5 on page 31, including how these were mitigated. These relate to the evaluation scope, challenges with national ethical clearance and the availability of evidence and results.
Question 10.	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 the principles in the 2020 revised UNEG Ethical Guidelines (integrity, accountability, respect, beneficence).	Yes	The report contains explicit and contextualized reference to the UNEG obligations of evaluators on page 29, in addition to WHO's ethical research guidance, UNICEF's Ethical Guidelines and standards for research and evaluation and Ethical Research Involving Children and UNICEF Guidance on Gender Integration in Evaluation.
	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.	Yes	There is a description of ethical safeguards for participants appropriate for the issues relevant to the methodology and how they are applied (see page 29). The evaluation involved interviews with children and is explicitly referring to UNICEF's procedures for Ethical Research Involving Children. In addition, ethical clearance and national and regional levels were obtained.
Question 11.	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 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.	Yes	The evaluation applied participatory methods, including establishing a Regional Adolescent Advisory Group composing of 8 girls and 3 boys from 7 countries between the ages of 15 and 19, of which one participant was visually impaired. Innovation practices were specifically requested in the ToR (under section 7 of the ToRs, point 3).
SECTION E:	EVALUATION FINDINGS (weight 25%)	67%	Comments on Rating
Question 12.	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.	Yes	The findings marshal sufficient levels of evidence to systematically address the evaluation questions, sub-questions and criteria. For example, all the questions are numbered and presented in the findings section of the report, and under each summary of the findings, the sources to the information on which the findings are based are specified, including references to documents, key informants and survey results. The questions and sub-questions are organized along the evaluation criteria.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	No	There is no explicit use of the intervention's results framework or Theory of Change in the formulation of the findings.
Question 13.	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 and quantitative data, but it does not present output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources, and it is stated (page 28) that: "The evaluation team triangulated findings from key informants both by ensuring similar questions were asked of all key informants and through the document review, ensuring different sources of data were validated, for example programme reports with external evaluations."
	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.	Yes	The findings presented are clearly supported by, and respond to, the evidence presented, including both positive and negative. For example, on page 32 it is stated that "The evaluation found that whilst all the case study countries appear to address girls' voice, agency and leadership to differing degrees as part of their adolescent-focused programming, this is not always done as part of a gender transformative programming approach", and on page 34: "we found strong evidence of programmes being designed to meet the needs of adolescents and young people based on robust analysis and research and secondary data, but across all case study countries, there was limited evidence of adolescents actually being involved in these processes". Findings are based on relevant types of evidence from the data collection processes.
	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 achievement or non-achievement of results are clearly identified. For example, on page 47, the reasons for why many UN agencies still work in silos and do not collaborate with others at the stages of programme planning or implementation are outlined in bullet points.
Question 14.	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.	Partially	While there was no assessment of the adequacy of the intervention's monitoring system to support decision making as such, the report states that UNICEF has put in place a system to track gender expenditure using both gender markers and activity level gender tagging. This is done at the country level and is led by the monitoring and evaluation officers. The evaluation team was unable to access this level of detailed information for the case study countries.
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	75%	Comments on Rating
Question 15.	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).	Yes	The conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking. For example, Conclusion 1 on page 79, which is based on findings 1.1, 1.2, 3.1, 9.1, 9.2 and 9.3, states that " Investing in adolescent girls' voice, leadership and agency. UNICEF ESAR Regional and Country Offices are investing in adolescent girls' voice, leadership and agency as a central approach to gender transformative programming, as laid out in the Gender Action Plan 3. This has led to positive changes among individual adolescent girls in relation to their confidence, self-esteem and self-belief. A significant focus on their agency and leadership skills appears to be more centred at the programme level, with less evidence emerging of how this has translated into changes affecting their lives outside of the programme, across the socio-ecological model. The ambition to be more transformative has been recently set out in the GAP 3, and therefore it is not surprising that there is limited understanding, and capacity, among staff and partners of how to implement more gender transformative programmes, which work across the socio-ecological model, and seek to change deep seated norms, and remove structural barriers. These strong outcomes at the individual level change are indicative of the focus of the previous GAP 2 and it is expected with continued emphasis, focus and resources girl-intentioned programming will over time become more transformative."
	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.	Yes	The conclusions are derived appropriately from findings, and provide clear reference to which findings each conclusion is drawn from. They present a picture of both the strengths and limitations of the intervention and they add insight and analysis beyond the findings. For example, Conclusion 2 on Sustainability, which is based on findings 4.1, 8.2 and 8.3, highlights both positive and negative findings, as follows: " UNICEF COs and Programmes are well aligned with government priorities, global frameworks, and UNICEF's strategic plans, as such several sustainable benefits can be found within UNICEF programming, including contributing to national policies and training curricula, ensuring they include a gender focus and address the rights of adolescent girls, some of which is contributing to transformative policy or legal change. However, more explicit emphasis is required to situate this work within the context of gender transformation, to address the root causes, and structural inequities among decision makers and lawmakers at all levels to engage with them in a more sustainable and gender transformative way."
Question 16.	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.	Partially	The identified lessons stem logically from the findings, and some of them have wider applicability and relevance beyond the object of evaluation (but not all).
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	The lessons are clearly presented, but they are not formulated as lessons. They are a mix of findings and recommendations. They do, however, have sufficient detail to be useful for the intended audience.
SECTION G: RECOMMENDATIONS (weight 15%)		88%	Comments on Rating
Question 17.	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.	Yes	The recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and conclusions. It is clearly specified (by way of numbering) which conclusion each recommendation is derived from. For example, Recommendation 1 derives from Conclusions 1, 3 and 4, and responds well to the evaluation purpose as the recommendations focus on how improvements to the programme may be implemented.
	ii Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Yes	The recommendations are useful and actionable for the primary intended users and uses, they are relevant to the intervention and appropriate guidance is given for the implementation. For example, it is stated under each recommendations who the primary intended users are and concrete guidance for implementation is given in bullet points under each recommendation.
	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 process for developing the recommendations is well described and it includes the active involvement of both duty bearers and rights-holders. Consultation workshops were held with the Evaluation Reference Group and the Regional Adolescent Advisory Group, where conclusions and recommendations were discussed to ensure they were practical, forward facing and grounded in the findings.
Question 18.	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.	Partially	There is a clear identification of the groups of duty bearers responsible for action for each recommendations. There is no clear prioritization or classification of the recommendations to support use, however.
SECTION H: REPORT STRUCTURE AND PRESENTATION (weight 5%)		50%	Comments on Rating
Question 19.	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.	Partially	The opening pages include the name of the evaluated object, date of the report, location of the evaluated object, names and organizations of the evaluators, name of the organization commissioning the evaluation, table of contents including figures, tables and annexes, list of acronyms and page numbers. Information about the timeframe of the evaluation is not included in the opening pages.
	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.	Partially	The annexes include the evaluation matrix, list of stakeholders consulted, list of documents consulted, data collection instruments, informed consent and other related documents. The annexes do not include a list of site visits (apart from an annex outlining the context of the 6 participating countries), or a document outlining the results chain/Theory of Change/logical framework. The Terms of Reference is missing; it is listed as Annex 14 in the table of contents, but it is not included in the Annex file.
Question 20.	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).	Partially	The structure is generally easy to identify and navigate, with numbered sections, clear titles and sub-sections. The report would benefit from better formatting, however, as it lacks a table of contents for the boxes, some annexes are missing (e.g. no ToR/Annex 14 is attached despite it being listed in the table of contents), etc.
	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.	Yes	The report structure accords to UNICEF guidelines for evaluation reports. The context, purpose and methodology precede the presentation of findings, which is followed by conclusions, lessons learned and recommendations.
Question 21.	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 report is not of a reasonable length; the main report is of 87 pages, while it was set out in the ToR/Inception Report (page 37) that the final report should be of maximum 50 pages excluding Annexes. The Annexes compose of another 197 pages.
	ii Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Partially	The report is generally easy to understand and written in an accessible way, but it does contain some sentences with grammatic errors and unclarities. (Example from page 5: " But this type of change takes time to demonstrate results at this level and since the approach has only recently been emphasised in the GAP 3, it is not surprising that there is not yet strong evidence of this yet."
	iii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	There is a frequent use of visual aids, with 17 figures and 8 tables. The report also contains boxes (example box 2 on page 17, but these are not clearly labelled and are not included in the table of contents.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	95%	Comments on Rating
Question 22.	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	While CRC, CCC or CEDAW are mentioned in the report, it is only in the context of UNICEF's work, and not in relation to the rights-based framework of this evaluation. The evaluation, however, did apply a strong rights-based framework in the design of the evaluation, with a written commitment to create a positive and safe environment for children and vulnerable adults who are connected to the evaluation work. This is outlined in detail in 11 points on pages 29-30.
	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	There is a clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation. In addition to a reference group, the evaluation established a Regional Adolescent Advisory Group composing of 8 girls and 3 boys from 7 countries between the ages of 15 and 19, of which one participant was visually impaired.
	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).	Yes	The language is empowering and inclusive, avoiding gender, heterosexual, age, cultural and religious bias. It uses the terminology of rights holders and duty bearers. Data is disaggregated by marginalized group and differential results are assessed whenever relevant or possible.
	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.	Yes	The evaluation assesses the extent to which the implementation of the intervention addresses child rights and Leave No-one Behind, and it is disability inclusive. For example, the survey demographic were disaggregated by disability in addition to gender, one of the adolescent participants in the advisory group was visually impaired and the evaluation reports under Evaluation Question 3 that "UNICEF engagement with people with disability through the programmes has still been limited and therefore there has been less alignment to the Convention on the Rights of Persons with Disabilities."

Question 23.	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	
	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 included an objective specific to assessment of human rights and gender equality considerations.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully integrated	The evaluation specified how gender issues were addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex. The evaluation methodology employed a mixed-methods approach, appropriate to evaluating GEWE considerations. A diverse range of data sources and processes were employed to guarantee inclusion, accuracy and credibility. The evaluation methods and sampling frame addressed the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate; and ethical standards were considered throughout the evaluation and all stakeholder groups were treated with integrity and respect for confidentiality.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Fully integrated	The evaluation has a background section that includes and intersectional analysis of the affected social groups; it specifies how data collection and analysis methods integrate gender considerations. The report provides specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area.

SWAP Rating Guidance

<p>i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.</p> <p>a. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results?</p> <p>b. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives?</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria?</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation?</p>
<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p> <p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex?</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations?</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility?</p> <p>d. Does the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate?</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality?</p>
<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p> <p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality?</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable?</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described?</p> <p>d. Does the evaluation report provide specific recommendations addressing GEWE issues, and priorities for action to improve GEWE or the intervention or future initiatives in this area?</p>