

GEROS Evaluation Quality Assurance Tool

Version: September 2021

Evaluation of the Child Labour: Exploitation of Children in South Asia Programme (2019 2024)

REPORT RATING SUMMARY		
Overall Rating	62%	Fair
●●●●● Exceptional (96% - 100%)	5	
●●●● Highly Satisfactory (87.5% - 95.99%)	4	
●●● Satisfactory (62.5% - 87.49%)	3	
●● - Fair (35% - 62.49%)	2	Meets UNICEF/UNEG standards for evaluation reports in some regards, but not all. Decision makers may use the evaluation with caution. Substantive improvements in some areas are needed
● - - Unsatisfactory (0% - 34.99%)	1	
REPORT DETAILS		
Title of the evaluation report	Evaluation of the Child Labour: Exploitation of Children in South Asia Programme (2019 2024)	
Report sequence number	ROSA, Nepal-64-2023-18894	
Region	SAR	
Year of report	2024	
Office	UNICEF Regional Office for South Asia	
Coverage (countries)	Region and Countries: South Asia, Pakistan, India	
ToRs present	Yes	
Date of review (dd/mmm/yyyy)	8.9.12 February 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	No	
Every child learns	No	
Every child is protected from violence and exploitation	Yes	
Every child lives in a safe and clean environment	No	
Every child has an equitable chance in life	No	
Gender equality (cross-cutting)	Yes	
Humanitarian action (cross-cutting)	No	
Evaluation object	Programme	
Evaluation type	Summative and formative	
Evaluation strategy	Mixed methods	
Evaluation design (primary method used)	Theory-based	
Evaluation level	Output & Outcome	
Geographic scope	Multi-country	
Primary SDG(s) covered (number)	5, 8, 16	
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>		
Key strengths of the evaluation		
<p>--The report is well structured (although there are some improvements to ensure accurate numbering, as suggested below).</p> <p>--Overall, the report is clearly written. However, as noted below, there are instances where language could have been revised for increased clarity.</p> <p>--The evaluation methodology received ethics review approval from the HML Institutional Review Board.</p> <p>-- The findings section is detailed and contains rich insights based on evidence collected and reviewed during the evaluation. However, as mentioned below, the report does not always list the source of the information (either by referencing it/them within the text or as footnotes) and this makes it difficult for the reader to understand what evidence is being used to create the finding. Evidence that supports both positive and negative findings are included in the findings section.</p> <p>--The evaluation uses credible forms of qualitative and quantitative evidence including programme documentation and reporting and opinions/recall of different key stakeholder groups (e.g., government, UN, civil society). However, there is a disconnect between the evaluation findings section and the information presented in the evaluation matrix.</p> <p>--The recommendations clearly link back to specific findings. This structure is useful in that it allows the reader to easily reference the corresponding finding(s) and understand the supporting evidence used in formulating each recommendation. This structure also clearly ties the recommendations back to the evaluation purpose (i.e. assessing specific OECD DAC criteria). Recommended actions under each recommendation are clearly directed to specific entities and prioritisation (temporal) is clearly indicated as well.</p> <p>--The annexes are appropriate and useful (although, as noted below, the numbering of these annexes requires revision for accuracy).</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>		
Areas for improvement		
<p>--Context: The report would have benefitted from a more detailed account of the status and needs of the rightsholders (children) in India and Pakistan, beyond child labour rates. For example, other indicators such as poverty, education, violence against children, etc. could have helped to build a more comprehensive picture of children's lived experiences.</p> <p>--The ToR, inception report and evaluation matrix contain a question regarding the ToC ('To what extent were the theory of change, assumptions, and linkages adequate to address the intended goals?') However, the ToC is not assessed within the final report. It is unclear why this was omitted.</p> <p>--The findings section requires a thorough review to ensure that the sources of information are cited in footnotes or explained in the body of the report (to better indicate that findings are based on triangulation and to allow readers to know the sources). Further, the report requires revision to align the EQs and sub-EQs presented in the findings section and the EQs, sub-EQs and questions from the ToR presented in the evaluation matrix.</p> <p>--The report is missing conclusions. This is a key component of any final report and should have been included.</p> <p>--Section 5 of the report jumps from 5.1 to 5.3. Section 4. Evaluation design, methodology and ethics has subsections that are labelled 3.5, 3.6, etc. and these require revision. The report numbering should have been revised to reflect final subheadings. Similarly, the appendices numbering requires review and revision (e.g., both the list of documentation and the evaluation matrix are titled Annex 2).</p> <p>--The report should have ensured that all images and boxes were titled. The report should have also ensured that all tables were readable, in the same format as other tables (e.g., font and text size) and not included as images.</p> <p>--The report should have included a list of interviewees/site visits in the appendices.</p>		

SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	50%	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).	Partially	The executive summary contains sufficient information to serve as a standalone document useful for informing decisionmaking. However, while the ToR does not indicate a page limit for the executive summary, it is longer than what is considered standard at 6 pages. It could have displayed recommendations in table format as well as statistics offered in the context section as a short table to meet the 5-page maximum set out in this criterion.
	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	The ToR does not indicate content for the executive summary. The executive summary includes an overview of the context and programme, evaluation objectives and methodology, evaluation audience, findings, lessons learned and recommendations. As noted elsewhere in this review, the final report lacks a conclusions section although they are required per the ToR and inception report.
	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	The executive summary provides a high-level summary of the context, programme, findings, lessons and recommendations. The executive summary does not introduce new information from what is presented in the rest of the report. The executive summary would have benefitted from additional information on the programme, specifically who the main partners were (duty bearers) and who the rightsholders were (children). The context of the intervention section of the executive summary could have been more concise to accommodate additional information on the programme itself. The executive summary does a good job at summarising findings per OECD DAC criteria and summarising the many recommendations from the body of the report.
SECTION B:	BACKGROUND (weight 5%)	43%	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.	Partially	The report contains a clear and relevant description of the programme including locations in India and Pakistan, timelines (2019-2023) and implementation status by indicating that it was a four-year programme that operated between 2019 and 2023. The report does not include any information around cost or budget for the programme.
	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	Table 2 indicates key programme stakeholders including rightsholders and duty bearers. This table also gives an overview of each stakeholders' roles. The report includes detailed information regarding the key stakeholders in India vs the key stakeholders in Pakistan (e.g., which specific government partners) within the paragraph preceding table 2. It would have been beneficial for table 2 to clearly indicate whether stakeholders were in Pakistan, India or both. The report does not indicate the numbers reached with disaggregation by gender, age, disability or other characteristics. Limitations of the MEL framework and available data regarding disaggregated data are noted within 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.	Partially	The report provides a clear and relevant description of the context of the programme with regard to relevant policy and international factors. However, the report would have benefitted from including reference to root causes of child labour which are tied to socio-economic contexts, etc. This insight would have been valuable to highlight in addition to the policy landscape.
	ii Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Partially	The SDGs are not mentioned in the context section of the report (titled background). Finding 1 links the programme's relevance to the SDGs. Finding 4, box 1 also links the programme to the SDGs. The report lists the following: 'These include: Target 5.3, which aims to eliminate all harmful practices; Target 8.7, which focuses on the eradication of forced labour and the worst forms of child labour; and Target 16.2, dedicated to ending abuse, exploitation, trafficking, and all forms of violence and torture against children.'
	iii Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	No	The report lacks an analysis of the status and needs of rightsholders beyond sharing statistics related to child labour rates. The report would have benefitted from including other indicators such as poverty rates, malnutrition, education, violence against children, etc. to give a fuller picture of the lived experiences of children affected by child labour in India and Pakistan.
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).	Partially	Table 2 contains a stakeholder map. However, the linkages between the different actors are unclear. Table 2 uses the titles 'duty bearers' for implementing agency(ies) which causes confusion. The table also lists 'CSO project implementing partners' but does not define what this means. Table 2 would have benefitted from some revision to more clearly map the actors and the linkages between them.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	Table 2 indicates the role of each mapped stakeholder. However, without additional details, it is unclear what some roles entail. For example, the organisations providing technical support had the role of 'support rolling out of the CPCMSS and CLS' but the report does not define or mention CPCMSS elsewhere in the report (this could have been a typo and the authors could have meant CPCMRS - this should still have been defined in the table itself for clarity). CLS is defined later on in the report, but table 2 should be clear regarding the acronym upon first usage.
SECTION C:		EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	42%	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	Section 3.1 clearly indicates the evaluation's purpose and objectives. The report indicates that the lessons learned, good practices and recommendations will be used for future programmes. Section 3.4 details the primary and secondary audiences of the evaluation. The report does not indicate why the evaluation was carried out at that point in time.
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).	Yes	The report clearly indicates the evaluation's four objectives. The final report does not note any changes made to the objectives included in the ToR. The QA reviewer confirms that the objectives of the ToR are the same as those presented in the final report.
	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).	Yes	Table 6 clearly outlines the temporal, geographic and thematic scope of the evaluation. However, the report states that programme implementation began in May 2019. The ToR states that the evaluation period begins in April 2019. This discrepancy should have been addressed/justified for clarity. While the report is clear regarding the temporal, geographic and thematic scope of the evaluation, thus fully meeting this criterion, the report would have benefitted from including the period under evaluation on the title page as this differs from the evaluation period and the programme implementation period. For example, the title page could have stated the following, noting that the discrepancy between April 2019 and May 2019 should be remedied: Programme implementation period: May 2019 - March 2024 Evaluation implementation period: July 2023 - January 2024 Period covered by the evaluation: April 2019 - May 2023
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.	No	The final report contains brief descriptions of the three outcomes (i.e., evidence generation, child protection system strengthening, establishing/reactivating community-based child protection systems), but does not contain a description of the programme's intended results or a clear description of the parts of the results chain that are applicable to/being tested by the evaluation. The full theory of change is available in annex 1, but is not discussed in the body of the final report.
	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 final report does not contain this information. A visual of the theory of change is available as a separate document in Annex 1. The report would have benefitted from including a simplified visual of the theory of change or a narrative version of the theory of change.
	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.	No	The report notes challenges in reconstructing the ToC (noted in section 2.1) but does not describe the reconstructed ToC nor provide a visual in the body of the final report. The findings do not assess the ToC, despite an evaluation question listed in the ToR (under relevance) relating specifically to the ToC: <i>'To what extent were the theory of change, assumptions, and linkages adequate to address the intended goals?'</i> This question is also listed within the evaluation matrix which is provided as an annex to the final report. However, an answer to this question is not provided in the final report.
SECTION D:		EVALUATION DESIGN AND METHODOLOGY (weight 20%)	73%	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.	Yes	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. Table 7 indicates the evaluation questions and corresponding OECD DAC criteria. The questions relate to the OECD DAC criteria. However, some questions listed under effectiveness could have been moved to sub-questions under EQ2 (which assesses the extent to which the programme achieved its intended outcomes and outputs). Annex 2 of the appendix (which should be titled annex 3 as the order is the ToC, list of documents consulted and then the evaluation matrix) contains the evaluation matrix which displays the evaluation sub-questions. These sub-questions respond to the evaluation's objectives including objective i (all sub-EQs), objective ii (all sub-EQs), objective iii (sub-EQs under EQ8) and objective iv (EQ1.3, 2.3)
	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 matrix includes indicators, methods, data sources and analysis type. The evaluation matrix does not include benchmarks or assumptions. The evaluation matrix requires a review as it is unclear how the questions from the ToR are addressed by the corresponding/referenced evaluation sub-questions. Further, the evaluation matrix lists EQ3 then EQ3.1 and then EQ5.2. It is unclear why there is a numbering error here.
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.	Partially	Most of the evaluation methods used (i.e., desk review, KIs and FGDs) are relevant and adequately robust for the evaluation's purpose, objectives and scope. These are clearly described in section 4, subsection 3.5 (mis-numbered as it should have been 4.5) That said, it is less clear why case file reviews were included as one of the methods. This method was not suggested in the ToR. The explanation within the inception report also does not provide sufficient justification for this method. The information in the final report is the same as that presented in the inception report and does not clearly justify the inclusion of the case file approach. As presented, it is unclear how the case file approach feeds into the evaluation.
	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 were appropriate and included both quantitative and qualitative sources (programme documentation and other secondary documentation, key stakeholders from government, CSOs, UNICEF, FCDO, INGOs (ILO and ActionAid), community influencers, child protection committee members, parents/families of children in child labour, children involved in child labour, among others). The data sources are clearly described in section 3.6 evaluation method (although this should have been titled 4.6).

iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or, if not, provide reasons for this).	Yes	<p>The final report clearly details the purpose sampling strategy for both site visits and stakeholder interviews/FGDs. The report includes the rationale for district selection in table 8. The report mentions that no clear document outlining the roles and responsibilities of various stakeholders in the programme was available during the inception stage. Consequently, the sampling methodology for KIIs and FGDs was formulated based using a UNICEF-compiled stakeholder mapping/identification. The report, however, does not comment on whether this was a limitation of the evaluation.</p> <p>The report included the selection criteria for KIIs and FGDs (pasted below for easy reference). The selection criteria included diverse perspectives by encompassing diverse geographic regions, gender balance, inclusion of marginalized groups and representation of different sectors.</p> <p>The report details the selection criteria as: <i>'the selection of participants for KIIs and FGDs was guided by the following selection criteria: (i.) Relevance to the evaluation objectives; (ii.) Experience or expert knowledge related to CLECSAP; (iii.) Representation of diverse geographic regions; (iv.) A balance of genders, (v.) Inclusion of marginalized groups and (vi.) Representation of different sectors, such as government, NGOs, and the community, rights holders, etc.'</i> The report states that <i>'this approach ensured a holistic and nuanced understanding of the programme's impact, addressing different viewpoints and experiences.'</i></p>
iv	Clear and complete description of the methods of analysis.	Partially	<p>The report would have benefitted from including information on why the data was coded per ToC assumptions. Further, the report would have benefitted from more clearly explaining how data was triangulated. As presented, the report states: <i>'Finally, data triangulation focused on different intervention strategies, countries, and states, district of intervention, implementation modalities (government, UNICEF, UNICEF's implementing partners, community members), comparing across different data collection methods to identify patterns and ensure cross-variable consistency.'</i></p>
v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	<p>The methodology described in the final report allows for drawing causal connections between outputs and outcomes. Specifically, EQ2 asks about outputs and outcomes, and data collected on that question can be used to draw causal connections.</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).	Partially	<p>The final report includes a list of possible challenges that were detailed in the inception report (tables 15 and 16). The final report notes that <i>'the inception report also included a list of potential limitations and key constraints that the evaluation may face, with regards to programme design (see Table 15) and data collection (see Table 16), along with mitigation strategies. Most of these constraints were particularly relevant for India.'</i></p> <p>The report would have benefitted from providing actual gaps/limitations and mitigation measures instead of potential gaps/limitations and potential mitigation strategies that were made available during the inception period. As written, the report indicates that information in tables 15 and 16 was compiled during the inception period.</p> <p>The report features one paragraph which summarises the evaluation team's efforts to mitigate the limitations faced. However, more details regarding how specific/actual limitations were mitigated is required.</p>
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	<p>The final report explicitly references UNEG obligations of evaluators as well as the UNEG ethical guidelines. The evaluation obtained ethical clearance from HML Institutional Review Board which confirmed the evaluation's compliance with ethical principles and standards.</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.	Yes	<p>The final report fully meets this criterion. The body of the final report (section 3.10) summarises the evaluation ethics and ethical clearance process. Annexes 6 and 7 include details of the ethics review process including the standards, consent forms, protocols and ethical approval from HML Institutional Review Board.</p> <p>Note: The final report states that Annex 4 details the ethical clearance process. However, Annex 4 is the evaluation terms of reference. References to annex numbers should have been updated and accuracy ensured during the final copyedit and proofread of the report.</p>

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.	No	The report does not contain any information on innovative practices used. Common data collection, analysis and validation methods were utilised (e.g., desk review, FGDs, KIIs, content and thematic analysis techniques, and validation of findings/recommendations with key stakeholders).
SECTION E: EVALUATION FINDINGS (weight 25%)		58%	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.	Partially	<p>While some findings are supported by sufficient levels of evidence, misalignment between the findings and the questions, coupled with the questions a lack of detail, means the report does not systematically address each question with the appropriate level of evidence. This is further detailed below.</p> <p>Section 3.3 of the final report indicate that the evaluation assessed five of the six OECD DAC evaluation criteria: <u>relevance, coherence, effectiveness, efficiency and sustainability</u>. (Note that the report is clear that gender, a human rights-based approach and disability inclusion were considered as cross-cutting evaluation criteria.) <u>However, the report does not contain any findings regarding coherence</u>. It is unclear why this was omitted. Further, it is important to note that the ToR and inception report do not reference coherence, so it is possible this was mistakenly placed into the final report.</p> <p>While findings are numbered, there are many instances where one numbered finding contains multiple, distinct findings. It is unclear why the authors combined multiple findings into one numbered finding. For example, finding 4 under effectiveness comments on the achievement of targets/indicators while also discussing the limitations in capturing the programme's full impact on child labour, especially regarding qualitative outcomes, gender and equity considerations and the effectiveness of interventions. This finding should have been presented as two separate findings – one on the achievement of targets/indicators and one regarding the results framework's limitations in its design. A third finding within finding 4 would have been regarding unintended results.</p>
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Yes	The programme's ToC is utilised in some of the findings, especially regarding output and outcome indicators and related 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	<p>The evaluation uses credible forms of qualitative and quantitative evidence including programme documentation and reporting and opinions/recall of different key stakeholder groups (e.g., government, UN, civil society).</p> <p>The findings section presents both output-level and outcome-level data.</p> <p>Triangulation is difficult to discern for many of the findings. The report does not consistently cite references or sources of data. In failing to do so, the reader is unable to know what statements are supported by what kind of evidence (e.g., programme reporting/documentation, one stakeholder, multiple stakeholders, documentation and stakeholders, etc.). The findings section requires a thorough review to ensure that sources are cited in footnotes or explained in the body of the report.</p>
	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	<p>The findings are detailed and supported by evidence - although sometimes the sources of evidence are unclear. The findings section contains rich insights based on evidence collected and reviewed during the evaluation. As mentioned above, the report does not always list the source of the information (either by referencing it/them within the text or as footnotes) and this makes it difficult for the reader to understand what evidence is being used to create the finding.</p> <p>Evidence that supports both positive and negative findings are included in the findings section. The findings include an analysis of the planned results and output indicators per the revised results framework (table 17). Several evaluation questions and sub-questions explored the 'extent to which ____'. These evaluation questions and subquestions did not have clear corresponding performance indicators, standards or benchmarks. However, the findings provided details regarding the extent to which the programme was able to progress towards a specific objective.</p>

	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).	Partially	Ofentimes, the report documents the achievements or non-achievements, but does not give the reasons for achievement/non-achievement. There are exceptions including in findings 11, 12, 13, 15, 16 among others. However, the report would have benefitted from investigating reasons behind achievement and non-achievement of results. This focus on the underpinning reasons would have helped to formulate detailed lessons learnt.
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	The final report contains some assessment of the adequacy of the MEL framework but, overall, the report does not contain an assessment of the programme's monitoring system to support decision-making. Box 1, for example, states that outcome and output indicators within the framework were specific and measurable with baselines and targets set for each. Box 1 also points out the shortcomings of the results framework (e.g., lack of qualitative indicators and outcomes for the children, lack of a full adoption of the leave no one behind principle in the MEL and data collection system).
SECTION F:	EVALUATION CONCLUSIONS & LESSONS LEARNED (weight 10%)	25%	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).	No	The report does not contain any conclusions. Conclusions were included as a requirement in both the ToR and inception report. These should have been included in the final report.
	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.	No	As mentioned above, conclusions were not included in the final report. This is a key component of any final report and should have been included.
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	Some lessons do not have supporting evidence from the evaluation (e.g., 1, 2, 4, 6, 7 and 8). It would have been helpful to summarise supporting evidence from the evaluation or to reference specific findings upon which the lessons were built. As phrased, the lessons have wider applicability and relevance beyond the object of the evaluation. However, the report does not provide supporting evidence for each lesson and it is unclear whether the <u>lessons are backed by findings from the evaluation</u> .
	ii Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Partially	Many of the lessons learnt do not contain sufficient detail to be useful for the intended audience. Many of the lessons learned do not describe situations where either positive or negative outcomes are connected to specific causal factors, and they don't elaborate on the circumstances in which this causality is valid. Each lesson should have been reviewed to ensure that each describe the situations in which the positive or negative lesson was learned and detail the causal factors.
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 evaluation purpose was to contribute to learning and accountability and to elicit lessons gained from the implementation of CLECSAP while evaluating its relevance, effectiveness, efficiency and sustainability (per the ToR). The recommendations section is structured in a way that links each recommendation to specific finding(s). This structure is useful in that it allows the reader to easily reference the corresponding finding(s) and understand the supporting evidence used in formulating each recommendation. This structure also clearly ties the recommendations back to the evaluation purpose (i.e. assessing specific OECD DAC criteria).
	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 were endorsed by the ERG and other major stakeholders through two validation workshops. As such, the QA reviewer should trust that these stakeholders endorsed recommendations that are useful and actionable for the intended users and uses. Upon a detailed review of the specific recommended actions, these appear to be both useful and actionable for the prescribed users and appropriate for the uses, and provide guidance for implementation.

	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).	Partially	The process for developing the recommendations is clearly described and it is clear that the development of recommendations included the involvement of duty-bearers. The report states that the recommendations received endorsements and were carefully examined and categorised based on their priority, the parties responsible for implementation and their relevance during two validation sessions. The report states that these validation sessions included the ERG and also involved 'all major stakeholders in both countries'. Without additional details regarding who attended the validation workshops, the QA reviewer cannot determine if both duty-bearers and rights holders were included. The report should have included more details regarding the stakeholder types present at the validation workshops and give an explanation for why rights holders were not involved, if applicable.
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.	Yes	The report clearly identifies the stakeholder(s) responsible for each recommendation and clearly classifies the each recommendation as short, medium or long term to support use. The report also prioritises some recommendations by stating the priority as 'high'. The remaining recommendations that are not deemed 'high' priority are not given a priority level (e.g., medium or low). The report would have been strengthened if each recommendation was given a priority level.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	57%	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 final report's opening pages include all of the listed elements in criterion 19.i except that it is lacking references to the boxes used throughout the report. As presented, only a list of tables is included. All boxes within the report should be numbered and titled. These should also be referenced within the table of contents section of the report (i.e., in the same manner as the list of tables).
	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 evaluation report should have listed the specific appendices, even if they are provided as a separate document. As displayed, the reader of the final report cannot see what is available in the appendices. A list of appendices with links or a table of contents for appendices alongside the note that it is available as a separate document would have been beneficial to include. In reviewing the separate document with appendices, the reviewer notes that the document contains many of the required annexes but is missing a list of interviewees and a list of site visits. Information on site visits is contained on page 5 of the report. The list of stakeholders interviewed (anonymised and disaggregated by sex, location and stakeholder type) should have been included. The report also includes other appropriate appendices such as the ethics review approval and ethical standards, consent forms and data collection protocols.
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 report is easy to identify and navigate as it uses numbered sections with clear titles and sub-titles. There are some numbering errors: --the body of the report states that the methodology is '4. Evaluation design, methodology and ethics'. However, the table of contents does not list this as section 4. Instead, section 4 is findings. --the numbering jumps from 5.1 Lessons Learned to 5.3 Recommendations. There is no 5.2 and this should have been addressed during a thorough copyedit/proofread of the final document.
	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 report structure largely follows the UNICEF guidelines for evaluation reports (context/background, purpose and methodology, findings, lessons learned and recommendations). The report is missing a conclusions section. Section 5 contains only lessons learned and recommendations. Conclusions should have been included in the final report.
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.	Yes	The evaluation ToR states that the page limit, if any, for the final report would be determined during the inception phase. The inception report does not include any page limitation or range for the final report. The language from the ToR makes clear that there will be no artificial limit for the final report length: 'Page limits, if any, to be established during the inception period. In general, there will not be artificial limits, but the report should aim for conciseness, readability, and visual appeal.' The report is 56 pages in length, exclusive of introductory information and the executive summary. Considering the guidance from the ToR quoted above, the report is of a reasonable length.
	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 would benefit from a thorough copyedit and proofread to improve readability. While a majority of the sentences are easy to understand, there are also many instances of run-on sentences or difficult-to-understand sentence structures. The report is largely free from grammar, spelling and punctuation errors.
	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 makes frequent uses of tables and boxes to help convey key information. These are clearly presented, labeled and referenced in the text with some exceptions: --there is an unnamed box on page viii of the executive summary --table 4 and 5 are presented as an image instead of a table. This means that the information conveyed in these tables is not readable to those using screenreaders and no alternate text is available. Further, the images are somewhat pixelated that makes it difficult to read. The font used does not match the rest of the tables. --there is a box in section 3.6.2.2 preceding table 9 that is not titled. --there is one box outlined under finding 9 that is not titled. --table 17 is presented as an image instead of a table. This means that the information conveyed in this table is not readable to those using screenreaders and no alternate text is available. Further, the image is somewhat pixelated that makes it difficult to read and the font does not match the rest of the tables. Because an image was used, the word 'labour' is underlined in red due to Word noting 'labour' as a misspelled word. --photos are not labelled and do not give information regarding copyright/ownership/location, etc.
SECTION I:	EVALUATION PRINCIPLES (weight 10%)	74%	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.	Partially	The report does reference and use some rights-based frameworks in the design of the evaluation. The report makes specific reference to the UNEG's Integrating Human Rights and Gender Equality in Evaluation as well as other guidances including the UNICEF Procedure for Ethical Standards in Research, Evaluations and Data Collection and Analysis, UNICEF Guidance on Gender Integration in Evaluation, and the UNICEF Procedure on Ethical Standards in Research, Evaluation and Data Collection and Analysis. How some components of these frameworks were used in the design of the evaluation is less clear. For example: --it is unclear how the evaluation met the checklist for a human rights and gender equality evaluation process as outlined in the UNEG HRGE in Evaluation guidance. E.g., Did the rapid evaluability assessment integrate the evaluability of human rights and gender equality dimensions of the programme? Were all stakeholder groups consulted at the end of the data collection stage to discuss findings and hear their views on the conclusions (not present in the final report) and recommendations of the evaluation? --it is unclear how the standards laid out in the UNICEF Guidance on Gender Integration in Evaluation were employed in the evaluation. For example, it appears that many standards outlined in Table D of the guidance were not considered.

	<p>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).</p>	Partially	<p>Table 2 provides an overview of duty bearers and rights holders. The report is clear regarding what stakeholder types were included as informants during data collection. However, the report is lacking a description around the level of participation of key rights holders and duty bearers in the conduct of the evaluation. The report references an evaluation reference group, but does not give details regarding the makeup of this group.</p> <p>The report states that the recommendations received endorsement from the members of the Evaluation Reference Group and 'other key project stakeholders during two validation sessions held within the evaluation timeframe – validation workshops carried out in Pakistan in October 2023 and in India in February 2023 – involving all major stakeholders in both countries.' Without further detail, the report lacks a clear description of the level of participation of key rights holders and duty bearers in the conduct of the evaluation.</p>
	<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>	Yes	<p>The report utilises empowering and inclusive language (e.g., children with disabilities, duty bearers, rights holders).</p> <p>Where available, the report provides disaggregated data (e.g., by male vs female). The report is clear regarding limitations around the availability of disaggregated data by other marginalised groups.</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>	Partially	<p>The report makes one reference to leave no one behind in box 1. The report simply states that the programme actively applied the leave no one behind principle, but then states that the MEL and data collection system did not explicitly mention gender or other equity considerations and that while some data was disaggregated by sex, key demographic details like age, migration status and disability were not integrated into the MEL framework or consistently analysed or reported. In this sense, the report does assess the extent to which the implementation of the programme addressed some components of the leave no one behind principle. However, it is unclear how the evaluation concluded that the programme actively applied the leave no one behind principle while also not integrating that principle into the MEL framework.</p>
<p>Question 23.</p>	<p>Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)</p>	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, the evaluation assesses whether sufficient information was collected and determined that this was a weak point of the programme.</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? Yes, the evaluation contained a specific objective: 'To assess the extent to which cross-cutting issues of gender, equity, disability, and human rights were integrated into the programme interventions.'</p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? Yes, the report cites a criterion titled 'Gender Equality, HRBA, Equity' which is addressed by EQ8.</p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? Yes, EQ8.</p>

ii	A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Fully 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? Yes, the methodology section states that the selection of participants for KIIs and FGDs was guided by selection criteria which included 'a balance of genders'. Further, the data collection instruments included administrative information about the interview which included the gender of the interviewee.</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? Yes, the evaluation employed a mixed-method approach.</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? Yes, the evaluation utilised document review/programme data and views of key stakeholders from different stakeholder groups to help guarantee inclusion, accuracy and credibility of findings.</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? The sampling was developed using UNICEF compiled stakeholder mapping and identification because there was no clear document highlighting the roles and responsibilities of numerous stakeholders in the programme at the evaluation's inception stage. The report states that the 'selection of participants for KIIs and FGDs was guided by the following selection criteria: (i.) Relevance to the evaluation objectives; (ii.) Experience or expert knowledge related to CLECSAP; (iii.) Representation of diverse geographic regions; (iv.) A</p>
iii	The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily 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? Somewhat. The report references the relative normative instruments and policies related to human rights, but does not include an intersectional analysis of the specific social groups affected by the issue.</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? Somewhat. The report is clear regarding limitations in the availability of disaggregated quantitative data. As noted above, it is often difficult to discern what types of stakeholders are informing findings.</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? The report offers achievements of unintended results on page 19. This includes issues around human rights and provides disaggregated data where available.</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? Yes, the report includes recommendations related specifically to GEWE issues (e.g., 'Strengthen gender lens in child protection and child labour programming', 'Support Social Welfare Department to address operational challenges of the DCPU, including in ensuring gender-sensitive child protection services and effective referral pathways', 'Intensifying efforts to inform the public about child protection issues and the critical importance of prevention. This</p>

SWAP Rating Guidance

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

- 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?