

# GEROS Evaluation Quality Assurance Tool

Version: September 2021

Multi-country evaluation of national childcare reforms with a focus on deinstitutionalization including CwD and difficult to place

## REPORT RATING SUMMARY

<b>Overall Rating</b>	78%	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. Decision makers may use the evaluation with confidence
●●○●●	Fair (35% - 62.49%)	2	
●○○●●	Unsatisfactory (0% - 34.99%)	1	

## REPORT DETAILS

<b>Title of the evaluation report</b>	Multi-country evaluation of national childcare reforms with a focus on deinstitutionalization including CwD and difficult to place
<b>Report sequence number</b>	ECARO_Switzerland/66/2024/18842
<b>Region</b>	ECAR
<b>Year of report</b>	2024
<b>Office</b>	UNICEF ECARO
<b>Coverage (countries)</b>	Bulgaria, Georgia, Moldova, Montenegro, North Macedonia, Serbia and Tajikistan
<b>ToRs present</b>	Yes
<b>Date of review (dd/mmm/yyyy)</b>	October 8, 2024
<b>Name of review firm</b>	IOD PARC

## CLASSIFICATION OF EVALUATION REPORT

<b>Management of evaluation (Managerial control and oversight of evaluation)</b>	UNICEF managed
<b>Unicef goal areas (Alignment with strategic plan priorities)</b>	
Every child survives and thrives	No
Every child learns	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
<b>Evaluation object</b>	Policy/Norms/Standards
<b>Evaluation type</b>	Summative and formative
<b>Evaluation strategy</b>	Mixed methods
<b>Evaluation design (primary method used)</b>	Theory-based
<b>Evaluation level</b>	Outcome
<b>Geographic scope</b>	Multi-country
<b>Primary SDG(s) covered (number)</b>	3, 10

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

### Strengths:

- The report is logically structured, well written and is accessible to a wide audience.
- The evaluation report provides visual aids throughout the report to help visualise key points.
- The report gathered data from a diverse stakeholder group, including persons with disabilities, giving the findings a rich data base.
- The report contains detailed contextual overviews for the countries and highlights key policies/conventions that are relevant to the object of the evaluation.
- The evaluation design and mixed-methods approach is well-described and appropriate for the evaluation's purpose and objectives.
- The report contains detailed ethical considerations.
- The report contains strong triangulation of qualitative and quantitative data to support findings, enhancing validity through cross-country comparisons. These findings offer rich insights for reflection and future efforts.
- The lessons learned are clear, applicable beyond the evaluation countries and are clearly linked to the evaluation evidence.
- Recommendations are logically derived from the findings and many include detailed guidance for implementation.

### Weaknesses:

- The report does not always accurately reference or title visuals (e.g., tables, charts, graphs).
- The evaluation report does not clearly note the intended uses of the evaluation.
- The report does not clarify why the evaluation was necessary at that specific time, which weakens the rationale for conducting the evaluation.
- Aside from recommendation 1, the recommendations do not specify the responsible party(ies) and are not prioritized or classified. The process for developing recommendations is not described.
- The annexes are missing a list of documents reviewed and the evaluation matrix.
- Evaluation sub-sections do not always clearly reference the specific evaluation question/sub-question being addressed, making it difficult to confirm whether each evaluation question and sub-question are adequately addressed.

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

- The report would have benefitted from a thorough copyedit/proofread to ensure that references to tables/figures are correct, table/figure numbering is accurate, and that references to tables in the body of the report are accurate.
- The executive summary should have included recommendations to be considered as useful for decision making.
- The stakeholder mapping should have been included in the report.
- The evaluation questions and sub-questions should have been organised in such a way that they align with the stated criteria.
- The findings section could have referenced the specific evaluation question/sub-question answered under each sub-section for clarity.
- Evaluation recommendations should have been prioritised and should include responsible party(ies) to carry out each recommendation.

## SECTION RATINGS

<b>SECTION A: EXECUTIVE SUMMARY (weight 5%)</b>	83%	Comments on Rating
<b>Question 1:</b> Can the executive summary inform decision-making?		
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 is clearly presented and meets the page requirements (maximum of five pages). The executive summary contains many of the elements required to serve as a standalone document but cannot be considered useful for informing decision making without more detailed recommendations. As written, the responsible party for each recommendation is not provided except for recommendation 1.

	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 necessary elements. The executive summary includes a very high-level overview of the intervention (i.e., 'national child care reforms undertaken between 2009 and the end of 2022, with a focus on deinstitutionalisation (DI) across the seven evaluation countries in Europe and Central Asia), the purpose and objectives, the intended audience, the evaluation methodology (and limitations), key findings per evaluation criteria and recommendations.
	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 evaluation does a good job at summarising the findings of the evaluation within five pages. It includes all significant information needed to understand the intervention. The summary findings allow the readers to have insight on the significant information to understand the interventions related to child care and DI reforms in the seven countries. The report does not introduce new information from what is presented in the rest of the report.
<b>SECTION B: BACKGROUND (weight 5%)</b>			71%	Comments on Rating
<b>Question 2.</b>		Is the object of the evaluation clearly described?		
	i	Clear and relevant description of the intervention, including: location(s), timelines, cost/budget, and implementation status.	Partially	There is a clear and relevant description of the intervention. Section 1.2 clearly lays out the description of the interventions including locations (national level), timelines/implementation statuses (noting the different national policy planning periods and UNICEF CPDs across the countries), and – to the extent that available data allowed – the cost/budgets associated with child care and DI reforms between 2009 and 2022 by country. That said, Table 1 includes financial data for years after 2022 (e.g., 2023-2027 in Bulgaria and Moldova, and 2023-2026 for Tajikistan) and it is unclear why these were included as these were clearly outside the temporal scope of the evaluation.  The final report details constraints around retrieving budgetary information related to child care and DI reforms.
	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/cities, sub-regions...) and in terms of numbers reached, with disaggregation by gender, age, disability... (as appropriate to the purpose of the evaluation).	Yes	Section 1.3 clearly describes the intended rights holders (all children under 18 except for children in conflict with the law who are remanded into custodial residential care by a court or given a custodial measure post-conviction) and duty bearers (by country, see Table 3: The main stakeholders and duty bearers across countries). As this evaluation deals with reform, actual numbers reached is not applicable. The section does, however, include the overall population of under-18s per country in Table 2.
<b>Question 3.</b>		Is the context of the intervention clearly described?		
	i	Clear and relevant description of the context of the intervention (i.e. relevant policy, socio-economic, political, cultural, power/privilege, institutional, international factors) and how context relates to the implementation of the intervention.	Yes	Section 1.1 provides a succinct description of the context of the intervention at a regional level, noting relevant UN conventions and policies of the EU. Section 2 details the regional and national contexts of the seven countries included under the evaluation (including demographic and socio-economic contexts, political contexts, and legal contexts).
	ii	Linkages drawn to the SDGs and relevant targets and indicators for the area being evaluated.	Yes	Section 1.6 details linkages to specific SDGs and targets (i.e., 1.3, 4.1, 5.1, 10.2, 10.3 and 16.6).
	iii	Clear and relevant description (where appropriate) of the status and needs of the rightsholders/beneficiaries of the intervention.	Yes	Section 2 of the final report provides a sufficient overview of the status and needs of the rights holders of the interventions. More detailed information is available in Section 2 of the inception report.
<b>Question 4.</b>		Are key stakeholders, their relationships and contributions clearly identified?		
	i	Identification of implementing agency(ies), development partners, right holders, and additional duty bearers and other stakeholders; and of linkages between them (e.g., stakeholder map) (if relevant).	No	The evaluation matrix used for country-level evaluations notes 'stakeholder mapping' as a data analysis technique, but the final report does not include stakeholder maps that identify implementing agencies, development partners, rights holders and additional duty bearers and other stakeholder and of the linkages between them. Stakeholder maps per country would have been valuable to include as annexes and refer to in the body of the final report.
	ii	Identification of the specific contributions and roles of key stakeholders (financial or otherwise), including UNICEF.	Partially	The evaluation report identifies the specific financial contributions of UNICEF in Table 1, but does not detail other financial contributions or roles of key stakeholders. The report would have been strengthened if it detailed the contributions per country and per stakeholder. However, the report indicates challenges in retrieving this data/whether this data is available:  "The evaluation scope includes all the government child care and DI reforms in the evaluation countries, irrespective of whether or not they were directly implemented with UNICEF support. Budgetary information for the reforms was gathered, though governmental financial data was limited across countries. Additionally, in many of the UNICEF CPDs, child care and DI components were integrated into various outcomes related to child protection, such as governance, social and behavioural change, education, health, and child rights monitoring, without any official breakdown of the amounts allocated directly or indirectly to child care and DI. Similar constraints related to external donor funding in some instances. These factors made identifying the total budget or cost for the evaluation's scope challenging. Nonetheless, summaries of the available budgetary and cost information across countries are provided in the table below. For more specific details on each budget see the country reports." (Pp. 2-3)
<b>SECTION C: EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)</b>			92%	Comments on Rating
<b>Question 5.</b>		Is the purpose of the evaluation clearly described?		

	i Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Partially	The purpose of the evaluation is clearly stated in Section 1.2 ('to assess the strengths and weaknesses in approaches taken by different countries; provide insight into system level bottlenecks and, in particular, to assess the results achieved by governments and UNICEF to date (outcome and impact level) in supporting children with disabilities and other highly marginalised and vulnerable children in the regions to remain with their families or in family-based care'). The final report does not clarify why the evaluation was needed at that point in time. Adding this detail would have strengthened the report.  Section 1.4 details the key intended users, but the evaluation report does not include intended uses of the evaluation. Intended uses should have been included.
<b>Question 6.</b>	Are the objectives and scope of the evaluation clear and realistic?		
	i Clear and complete description of what the evaluation seeks to achieve by the end of the process with reference to any changes made to the objectives included in the ToR (if applicable).	Yes	The purpose and four objectives of the evaluation are detailed in Section 1.2. The objectives listed in the evaluation report match those included in the ToR.
	ii Clear and relevant description of the scope of the evaluation: what will and will not be covered (thematically, chronologically, geographically with key terms defined), as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	Section 1.5 details the scope of the evaluation which aligns with what was presented in the ToR, including the thematic, geographical, and chronological scope. The evaluation is clear that rights holders for this evaluation do not include children in conflict with the law who are remanded into custodial residential care by a court or given a custodial measure post-conviction.  The report contains the same rationale as the ToR regarding the non-inclusion of programming in the education sector, health sector and social protections sector (due to resource constraints).
<b>Question 7.</b>	Is the theory of change, results chain or logic well articulated?		
	i Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	Section 2.4 contains a broad regional ToC developed during the inception report (Figure 3 followed by the ToC in narrative form). The evaluation notes that reconstructed ToCs for each country are available in the country evaluation reports.
	ii Causal relationship between outputs and outcomes is presented in narrative and/or graphic form (e.g., results chain, logic model, theory of change, evaluation matrix).	Yes	The causal relationships between outputs and outcomes are presented in both graphic form (Figure 3) and in narrative form within Section 2.4.
	iii For theory-based evaluations, the theory of change or results framework is assessed, and if requested in the ToR, it is reformulated/improved by the evaluators.	Yes	The ToC was reconstructed during the inception phase (both country-level ToCs as well as a broad, regional ToC). The reconstructed country-level ToCs are available in the country reports and the regional ToC is available in Section 2.4.
<b>SECTION D:</b>	<b>EVALUATION DESIGN AND METHODOLOGY (weight 20%)</b>	86%	Comments on Rating
<b>Question 8.</b>	Does the evaluation use questions and the relevant evaluation criteria that are explicitly justified as appropriate for the purpose of the evaluation? <i>UNICEF evaluation standards refer to the OECD/DAC criteria - Relevance; Coherence; Effectiveness; Efficiency; Sustainability; Impact (not all are necessarily relevant for all evaluations). Evaluations should also consider equity and leaving no-one behind, gender and human rights based approach (these can be mainstreamed into other criteria). Humanitarian evaluations should also consider Coverage; Connectedness; Coordination; Protection; Security.</i>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Partially	The evaluation questions vary between country-level evaluations and the regional synthesis report.  Country-level evaluation questions and sub-questions are appropriate for meeting the purpose and objectives of the evaluation, but do not specifically address all objectives. For example, objective 3 is only assessed in Question 26 (What is UNICEF's role in the sustainability of the DI reforms?) and is not explicitly addressed as a question or sub-question under effectiveness/impact.  The evaluation questions and sub-questions are organised by evaluation criteria, but the questions do not always align with the stated criteria. For example: -- Question 3 (To what extent, if any, have UN guidelines on alternative care for children, EU guidelines and policies, CRC and CPRD on the national DI reforms acted as a trigger for national reform?) is listed under 'Relevance', but this question and its sub-questions are appropriate for 'Effectiveness' (i.e., assessing whether guidelines/conventions 'triggerred' actions).  --Question 25 (Which key reform partners influenced / continue to influence the DI reforms implementation, performance and sustainability?) is listed under 'Sustainability' but is also asking about reform implementation and performance instead of focusing on the sustainability of efforts.  Similarly, the synthesis report key evaluation questions do not specifically address objective 3 (To assess the actual and potential contribution of UNICEF's work to the national progress (outcome and impact) in deinstitutionalisation child care reforms for children in residential care, including for children with disabilities and other 'difficult to place children). UNICEF's actual/potential contributions are not explicitly addressed as a question or sub-question under effectiveness/impact.

	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 report states that an evaluation matrix for the synthesis report was not used. As stated in the report: 'An evaluation matrix was not created for the regional report as reliance was placed on the individual country matrices.' It would have been useful to provide an evaluation matrix for the regional synthesis report.  The evaluation matrix for country-level evaluations includes questions and sub-questions per evaluation criteria alongside indicators/measures, data collection methods, data analysis techniques and data sources.
<b>Question 9.</b>		Does the report specify adequate methods for data collection, analysis, and sampling?		
	i	Evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope; and are fully and clearly described.	Yes	The evaluation design and set of methods are relevant and adequately robust for the evaluation's purpose, objectives and scope. These include desk review, and primary data collection with a diverse set of stakeholders via KIIs and FGDs. The design/set of methods are fully and clearly described in Section 3 of the final report.
	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 are appropriate, contain both quantitative and qualitative sources, and are clearly described in Section 3.3 and Figure 4.
	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 described in Section 3.4. The explanation includes a description of how diverse sites were chosen (i.e., capital city as well as locations with a high prevalence of residential facilities and child care service providers; mix of rural and urban areas). Section 3.4 also details the respondent selection as primarily purposive and non-random for all countries and prioritised diversity to ensure respondents of diverse backgrounds and with diverse perspectives were included in the evaluation. The sampling included key informants from regional, national and sub-national levels who had knowledge, expertise and accurate information in relation to child care and DI reforms, including stakeholders in relevant government agencies, UN agencies and international and national NGO partners. FGD participant sampling strategy for interviewing care-users (children, adolescents and adults) aimed to ensure representation across different disabilities. The final report states that identification of the sample relied on organisations that were involved in KIIs and FGDs, utilising a snowball sampling method. FGDs with other stakeholder groups were also chosen purposively to ensure inclusion of perspectives from local services providers, residential institutions and NGOs.
	iv	Clear and complete description of the methods of analysis.	Yes	Section 3.5 clearly and completely describes the data analysis methods. These include qualitative data analysis through MAXQDA (coding to identify key themes, patterns and relationships relevant to the evaluation questions), a realist evaluation approach to data analysis, participatory systems mapping and agent-based modelling, quantitative data analysis via Excel of administrative data (secondary data).
	v	Methodology allows for drawing causal connections between outputs and expected outcomes.	Yes	The mixed-methods approach and analysis techniques described in Section 3 allow for drawing causal connections between outputs and expected outcomes.
	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	Section 3.6 clearly and completely describes the limitations and mitigation measures of the evaluation. The mitigation measures address how gaps in evidence were addressed by the evaluators.
<b>Question 10.</b>		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	Section 3.8 explicitly references the UNEG obligations of evaluators as well as many of the principles in the 2020 revised UNEG Ethical Guidelines. These considerations/obligations are contextualised to this evaluation with the final report describing how each was approached during the 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	The ethical safeguards are provided in detail within the inception report and are summarised in the final report. The evaluation included children and Section 3.8 explicitly references the UNICEF procedures for Ethical Research Involving Children.  The evaluation underwent an ethical review process conducted by the HML Institutional Review Board (IRB) which comprehensively assessed the research protocol / methodology, data collection methods and plan and ethical protocol and tools, along with the data collection tools, information sheets and consent forms. The evaluation includes the ethical approval in Annex F with template consent forms available in Annex G.
<b>Question 11.</b>		Does the evaluation incorporate innovative practice that adds value to the evaluation process?		
	i	Does the evaluation incorporate innovative practice that adds value to the evaluation process?	Partially	The evaluation report indicates that it has sought to employ a theory-based approach that also draws on a realist approach alongside participatory systems mapping and agent-based modelling. While these approaches and methods are not entirely innovative (and are identified in the ToRs as 'appropriate designs for questions'), they demonstrate a thoughtful approach by the evaluation team. However, as there is limited description of these approaches and methods (and how they were applied) in the report, including the Annex, this is rated as 'partially' meeting this criteria.
<b>SECTION E:</b>		<b>EVALUATION FINDINGS (weight 25%)</b>	<b>83%</b>	<b>Comments on Rating</b>
<b>Question 12.</b>		Do the findings clearly address all evaluation objectives and scope?		

	i Findings marshal sufficient levels of evidence to systematically address all of the evaluation's questions, sub-questions and criteria.	Partially	The report excels in providing sufficient levels of evidence to address the evaluation criteria. The findings section also appears to coincide with many evaluation questions and sub-questions. However, the titles of the sub-sections make it difficult to see which evaluation question or sub-question is being answered in each sub-section (for example, see sub-sections in Section 5 compared to the evaluation questions under Effectiveness). To remedy this, the report could have referenced the specific question/sub-question for each sub-section.
	ii Explicit use of the intervention's results framework/ToC in the formulation of the findings.	Partially	The evaluation findings describe specific outputs and outcomes, implying the use of a results framework (or multiple results frameworks i.e., one per country). The evaluation findings also outlines causal factors that contribute to achieving or not achieving results which reflects the 'if-then' logic of the regional ToC presented in Figure 3 of the final report.  While the findings align with the principles of a results framework and reflect the logic used in the regional ToC, the findings do not explicitly articulate/reference the results frameworks/ToC in the findings section.
<b>Question 13.</b>	Are evaluation findings derived from the conscientious, explicit and judicious use of the best available, objective, reliable and valid data and by accurate quantitative and qualitative analysis of evidence.		
	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident through the use of multiple data sources.	Yes	Triangulation is evident within the findings. The evaluation uses credible forms of qualitative data (e.g., KII and FGD findings, secondary sources such as legal frameworks, policies, UNICEF sources, etc.) as well as quantitative data (e.g., administrative data). The findings also incorporate cross-country comparisons which strengthens the analysis by providing context and validating the findings through different country examples. The findings present both output (e.g., implementation of policies/legal frameworks, development of standards, establishment of gatekeeping mechanisms, number of children leaving residential care) and outcome level data (e.g., transition of children into mainstream education and reduction of children in institutional care, shifts in public attitudes/change in support for institutional care, impact of family support services) per evaluation criteria.
	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	Findings are clearly supported by/respond to the evidence presented and include both positive and negative findings. The findings are based on country-level ToC indicators which were translated to a regional-level ToC. The findings are organised by evaluation criteria and coincide with the evaluation questions (note that sometimes multiple evaluation questions are contained under one heading e.g., Section 4.1 addresses key evaluation questions 1 and 2 and there are other times - such as sub-sections under Section 5 which it is difficult to discern which evaluation question coincides to the sub-section).
	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	The findings do a good job at assessing the causal factors that led to achievement or non-achievement of outputs/outcomes (e.g., stigma and perceptions, limited capacity and skills among social welfare service providers, limited or lack of community-based services and family-based alternatives for children with disabilities, closure of large-scale institutions - leading to reduction in institutionalised care). The findings include analysis of the logical chain, focusing on what drove progression or lack of progression from implementation to results (e.g., how countries that prioritised family support and community-based services saw significant reductions in number of children in residential care).
<b>Question 14.</b>	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's monitoring system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Yes	The evaluation includes assessments of various monitoring systems that feed into tracking the effects of child care and DI reforms across the seven countries. For example, the evaluation notes the low level of monitoring of the placement of children post-reintegration, the lack of monitoring data on the outcomes of reintegration across the countries, inadequacy of systems of monitoring and oversight of accessibility for children with disabilities in schools, among others. Section 5.9 Effectiveness of systems for monitoring and reporting on DI reforms details the levels of attention given to reporting on the results of DI reforms across the evaluation countries.  As the evaluation was not based on one intervention but rather all child care and DI reforms across the countries between 2009 and 2022, the assessment of the completeness and appropriateness of results/performance framework – including vertical and horizontal logic, M&E tools and their usage to support decision making is not-applicable.
<b>SECTION F:</b>	<b>EVALUATION CONCLUSIONS &amp; LESSONS LEARNED (weight 10%)</b>	75%	Comments on Rating
<b>Question 15.</b>	Do the conclusions clearly present an objective overall assessment of the intervention?		
	i Conclusions are clearly formulated and reflect the purpose and objectives of the evaluation. They are sufficiently forward looking (if a formative evaluation or if the implementation is expected to continue or have additional phase).	Partially	The conclusions section (Section 9.1) is quite long and provides a comprehensive summary of findings rather than a true conclusions section. The conclusions section should have synthesised key insights drawn from the evaluation's findings and their implications, rather than simply restating what has been observed. That said, the conclusions section does contain some forward-looking language and provides direction for future actions.

	ii	Conclusions are derived appropriately from findings, and present a picture of the strengths and limitations of the intervention that adds insight and analysis beyond the findings.	Partially	The conclusions are derived from findings and present a picture of strengths and limitations of the interventions across the seven evaluation countries. However, the conclusions section does not consistently add insight and analysis beyond the findings. As noted above, there is some forward-looking language and the conclusions section does provide some direction for future actions.
<b>Question 16.</b>		Are logical and informative lessons learned identified? [N/A if lessons are not presented and not requested in ToR]		
	i	Identified lessons stem logically from the findings, have wider applicability and relevance beyond the object of the evaluation.	Yes	The four lessons learned are connected to the evidence. They have wider applicability and relevance beyond the seven countries, while reinforcing the key messages from the evaluation.
	ii	Lessons are clearly and concisely presented, yet have sufficient detail to be useful for intended audience.	Yes	The lessons are clearly presented but could be more concise (e.g., revised to reduce some of the longer, more complex sentences without losing their meaning). However, at only half a page in length, they are sufficiently concise for the evaluation report. The lessons have sufficient details to ensure that lessons offer valuable insights for the intended audiences and lead into the recommendations that follow.
<b>SECTION G:</b>		<b>RECOMMENDATIONS (weight 15%)</b>	<b>38%</b>	<b>Comments on Rating</b>
<b>Question 17.</b>		Are recommendations well grounded in the evaluation?		
	i	Recommendations align with the evaluation purpose, are clearly formulated and logically derived from the findings and/or conclusions.	Yes	The 8 recommendations and sub-recommendations/recommended actions align with the evaluation purpose of assessing the strengths and weaknesses in approaches taken by different countries and to assess the results achieved by governments and UNICEF to date in supporting children with disabilities and other highly marginalised and vulnerable children in the region to remain with their families or in family-based care. The recommendations logically derive from the findings and are clearly formulated.
	ii	Recommendations are useful and actionable for primary intended users and uses (relevant to the intervention); guidance is given for implementation, as appropriate.	Partially	Overall, the recommendations are useful and actionable. Aside from recommendation 1 which is intended for UNICEF, the recommendations do not specify the responsible parties. Each recommended action should specify the responsible party(ies). In order to determine whether the recommendations are actionable for the intended users, the users must be specified.  The authors provide specific guidance for implementation for several of the recommendations (e.g., recommendation 1, 2, 3, 4, 5, 7).
	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).	No	The report does not include information on the process for developing the recommendations and this should have been provided.
<b>Question 18.</b>		Are recommendations clearly presented?		
	i	Clear identification of groups or duty-bearers responsible for action for each recommendation (or clearly clustered group of recommendations). Clear prioritization and/or classification of recommendations to support use.	No	Aside from recommendation 1 which is directed at UNICEF, the recommendations do not identify specific groups or duty bearers responsible for action for each recommendation/clustered group of recommendation. The recommendations are not prioritised or classified to support use.
<b>SECTION H:</b>		<b>REPORT STRUCTURE AND PRESENTATION (weight 5%)</b>	<b>86%</b>	<b>Comments on Rating</b>
<b>Question 19.</b>		Does the evaluation report include all relevant information?		
	i	Opening pages include: Name of evaluated object, timeframe of the evaluation, date of report, location of evaluated object, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents -including, as relevant, tables, graphs, figures, annexes; list of acronyms/abbreviations, page numbers.	Yes	The opening pages include the name of evaluated object, timeframe of the evaluation, date of report (year), location of evaluated object (within Acknowledgements section), names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents, list of tales and list of figures alongside page numbers.
	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 terms of reference, list of participants, research sites, and data collection tools. The annexes also include the ethical protocol, Coram International's ethical guidelines for field research, ethical approval, and participant information sheets and consent forms.  The annexes are missing the list of documentary evidence. An evaluation matrix for the synthesis report was not created but drew on the country-level evaluation matrix. These are available in the inception report but should have also been included as annexes to the final report.
<b>Question 20.</b>		Is the report logically structured?		
	i	Structure is easy to identify and navigate (for instance, with numbered sections, clear titles and sub-titles, well formatted).	Yes	The report is easy to identify and navigate. It uses numbered sections with clear titles and sub-titles. The report is well formatted and uses hyperlinks within the table of contents to facilitate easy navigation.
	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 final report structure accords to UNICEF guidelines for evaluation reports. Background/context and purpose of the evaluation is presented first, followed by the methodology and then evaluation findings, conclusions and lessons learned and recommendations.
<b>Question 21.</b>		Is the report well presented?		
	i	Report is of reasonable length; it does not exceed number of pages that may be specified in ToR.	Yes	The final report is of a reasonable length and does not exceed the page limit of 60 pages per the ToR.
	ii	Report is easy to understand (written in accessible way for intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	The report is easy to understand and general free from grammar, spelling and punctuation errors. A final copyedit/proofread may have caught the remaining grammatical/punctuation errors, but these are few and far between and do not impact the reader's ability to consume the information.

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	<p>The report makes frequent use of visual aids which convey key information and add richness to the report. Further, there is a list of figures and list of tables following the acronyms list for easy reference.</p> <p>Nearly all of the visuals are clearly labelled. However, there are some exceptions to this. The following items lack a title but are still referred to in the text:</p> <ul style="list-style-type: none"> <li>--Figure/table in Section 1.6 which presents the SDG linkages</li> <li>--Table that discusses the relationship to the EU under Section 4.2</li> <li>--The chart under section 5.3 which presents old-style institutions and new small-scale residential care</li> </ul> <p>There are also some visuals/figures that are titled but are not specifically referenced by figure/table number in the text. The information relayed in the visuals is, however, discussed in the surrounding text.</p> <ul style="list-style-type: none"> <li>--Table 2: Number of children across countries</li> <li>--Figure 1: Map of countries covered in the report</li> <li>--Table 4: Demographic and Economic Data for countries involved in report</li> <li>--Figure 2: Timeline of most significant political events across countries</li> <li>--Figure 3: Reconstructed Theory of Change</li> <li>--Figure 4: Snapshot of the methodological approaches use</li> <li>--Table 6: Overview of main conventions/agreement/guidelines and the year they were ratified</li> </ul> <p>There are also some instances where figures are referred to incorrectly:</p> <ul style="list-style-type: none"> <li>--Section 5.1 states 'As demonstrated in table 3 below', but the visual is actually titled 'Figure 5'.</li> <li>--Section 5.1 states 'As set out in table 3' but the table is actually titled 'Table 8'.</li> <li>--Section 5.6 states 'as is demonstrated by table 4 below' but the table is actually titled 'Table 9'.</li> </ul> <p>In Section 3.6, a table but stands alone as the entire risks and limitations section. An introductory sentence or two prior to the table would have helped the document flow better, but it is not required</p>
<b>SECTION I: EVALUATION PRINCIPLES (weight 10%)</b>		<b>100%</b>	<b>Comments on Rating</b>
<b>Question 22.</b> Did the evaluation design and style consider incorporation of the UN and UNICEF's commitment to a human rights-based approach to programming, to gender equality, and to equity?			
i	Reference and use of rights-based framework, and/or CRC, CCC, CEDAW and/or other rights related benchmarks in the design of the evaluation.	Yes	The report references and makes use of several rights-based frameworks including the CRPD, CRC, EU Strategy for the Rights of Persons with Disabilities 2021-2030, EU Strategy on the Rights of the Child 2021-2024 and the Council of Europe has also issued a Disability Strategy 2017-2023.
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	The report describes the level of participation of key rights holders and duty bearers as informants. The evaluation also had an evaluation reference group (ERG) whose roles and responsibilities are described in section 3.7. The ERG comprised representatives from the relevant line ministries, institutions/service providers, CSOs, and independent bodies. Additionally, the report describes the roles of UNICEF ECARO and country offices as key actors for validation of evaluation deliverables and oversight.
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 report uses language that is empowering and inclusive. It avoids gender, heterosexual, age, cultural and religious bias. The report uses the terminology of rights holders and duty bearers (but also references beneficiaries). Where data is available, the evaluation shows disaggregated data by different groups e.g., children with disabilities. The differential results of some marginalised groups such as children with disabilities is referenced.
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 interventions address child rights and Leave No One Behind. The evaluation includes disability within the context section, within the evaluation questions and within the evaluation findings and recommendations.
<b>Question 23.</b> Does the evaluation meet UN SWAP evaluation performance indicators? (Note: this question will be rated according to UN SWAP standards with detail provided below)		9	
i	GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	<p><b>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.</b></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? <b>YES.</b></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? <b>YES (MAINSTREAMED).</b></p> <p>c. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? <b>MAINSTREAMED.</b></p> <p>d. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? <b>YES.</b> (Question 8.)  <input type="checkbox"/></p>

	<p>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</p>	<p>Fully integrated</p>	<p><b>SWAP Rating Guidance: ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</b></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? <b>YES.</b></p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEWE considerations? <b>YES.</b></p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? <b>YES.</b></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? <b>YES.</b></p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? <b>YES.</b></p>
	<p>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</p>	<p>Fully integrated</p>	<p><b>SWAP Rating Guidance: iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</b></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? <b>YES.</b></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? <b>YES.</b></p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? <b>YES.</b></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? <b>NO.</b></p>

**SWAP Rating Guidance**

<p><b>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.</b></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><b>ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.</b></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><b>iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.</b></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>